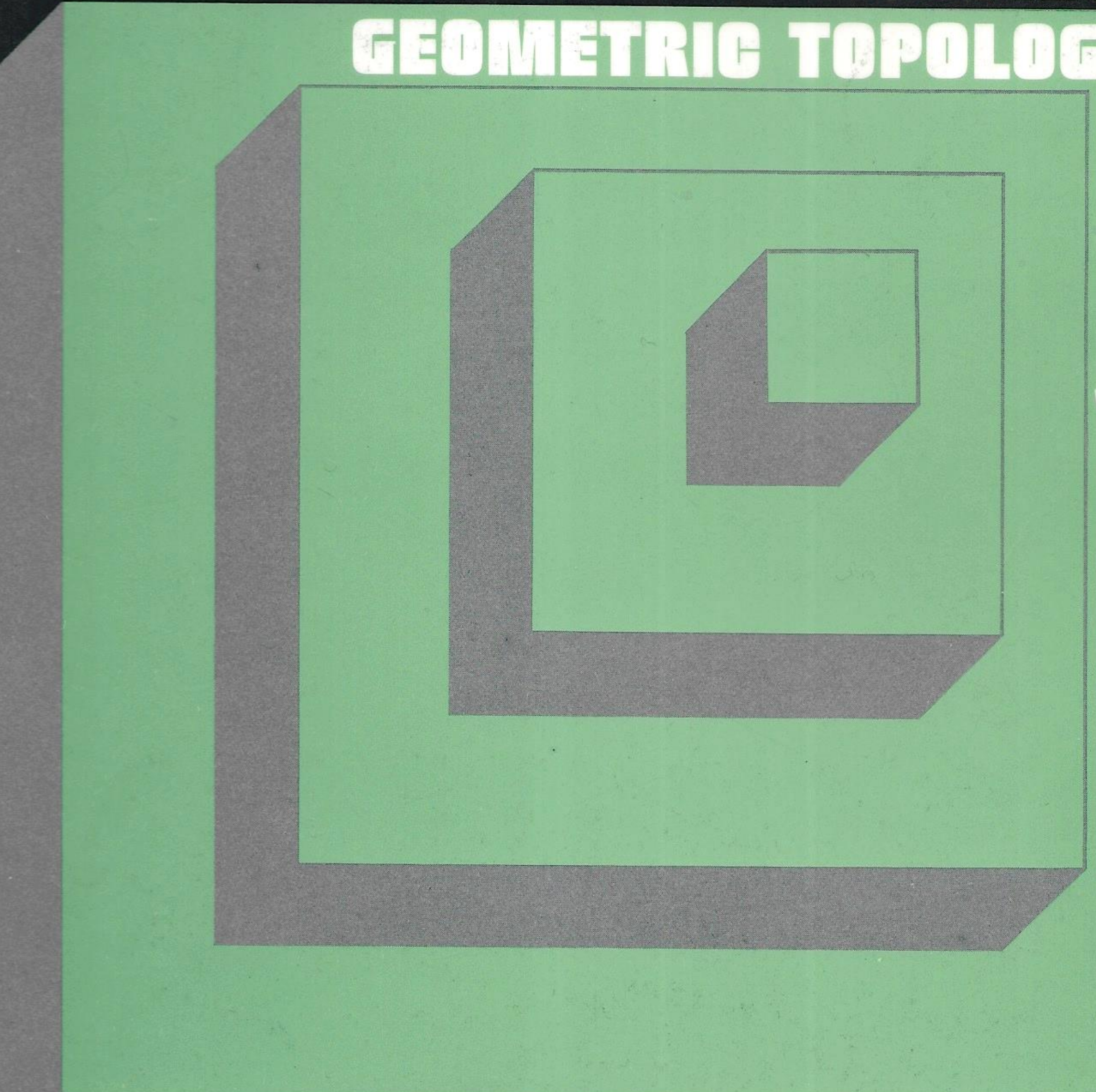
M335 GT 7, 8

THE OPEN UNIVERSITY
Mathematics: A Third Level Course
M335 Studies in Pure Mathematics

STUDIES IN PURE MATHEMATICS



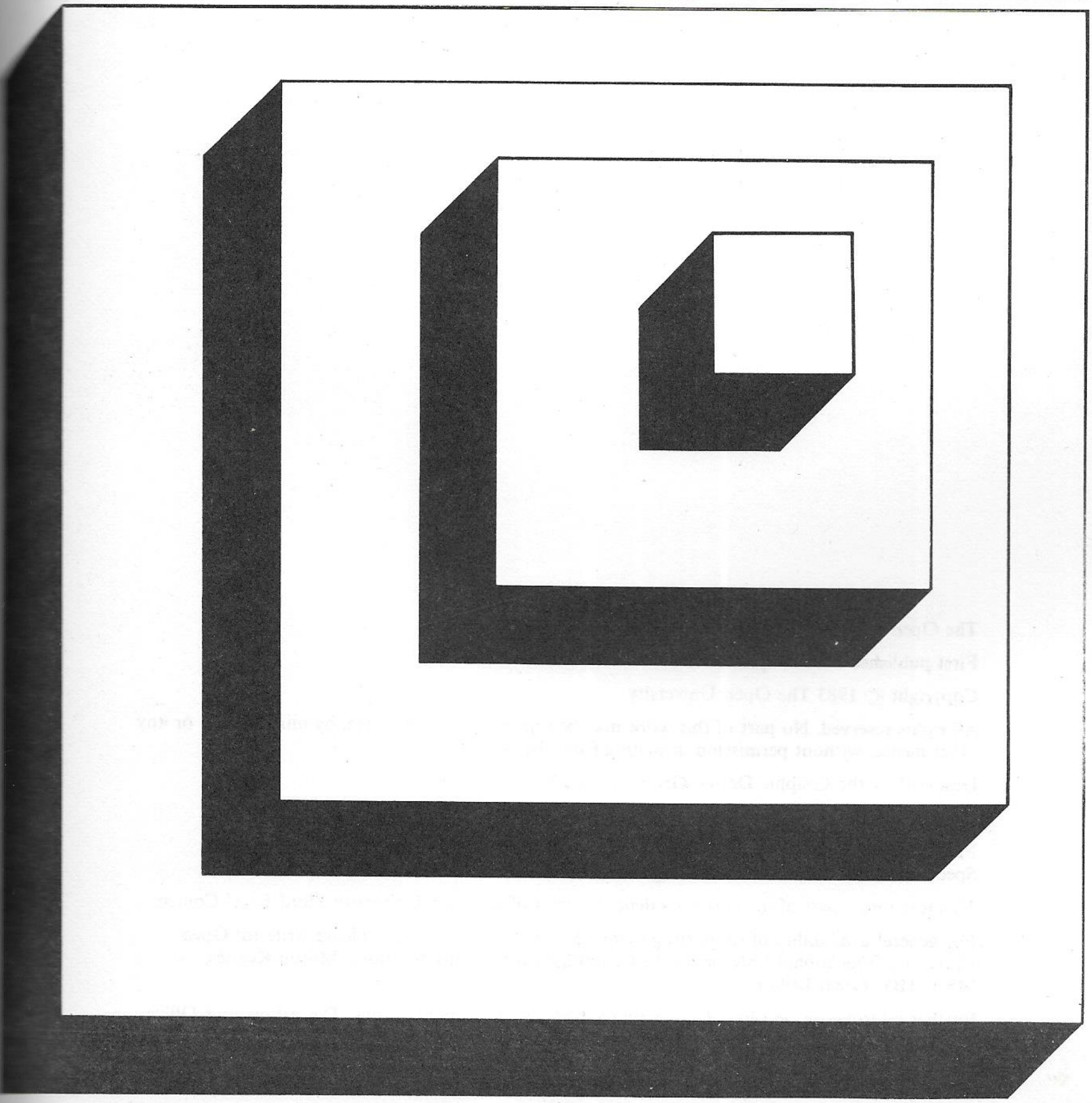
GEOMETRIC TOPOLOGY

UNITS 7,8

THE OPEN UNIVERSITY
Mathematics: A Third Level Course
M335 Studies in Pure Mathematics
Prepared by the Course Team

GEOMETRIC TOPOLOGY

UNITS 7.8



The Open University Press, Walton Hall, Milton Keynes.

First published 1983. Reprinted 1984.

Copyright © 1983 The Open University

All rights reserved. No part of this work may be reproduced in any form, by mimeograph or any other means, without permission in writing from the publishers.

Designed by the Graphic Design Group of the Open University.

ISBN 0 335 14021 1

Produced in Great Britain by
Speedlith Photo Litho Limited, Longford Trading Estate, Manchester M32 0JT.

This text forms part of the correspondence element of an Open University Third Level Course.

For general availability of supporting material referred to in this text, please write to: Open University Educational Enterprises, 12 Cofferidge Close, Stony Stratford, Milton Keynes, MK11 1BY, Great Britain.

Further information on Open University courses may be obtained from: The Admissions Office, The Open University, P.O. Box 48, Milton Keynes, MK7 6AB.

1.2

Contents

Unit 7 Winding Numbers and Brouwer Degree

7.1 Winding Number **6**

7.2 A Formula for the Winding Number **17**

7.3 Proof of the Brouwer Degree Theorem **20**

7.4 Applications of Winding Number **27**

Appendix I **30**

Appendix II **31**

Unit 8 Poincaré Index Theorem; Revision Module

8.0 Introduction **34**

8.1 Singularities of Flow Fields **34**

8.2 Flows on Surfaces **41**

8.3 Revision of the Course **48**

UNIT 7

WINDING NUMBERS AND BROUWER DEGREE

7.1 Winding Number

This section, which includes TV and TV-related Audio, deals with a rigorous definition of *winding number* in terms of covering spaces. It concludes with an important application of winding number — the *Brouwer degree* of a continuous function.

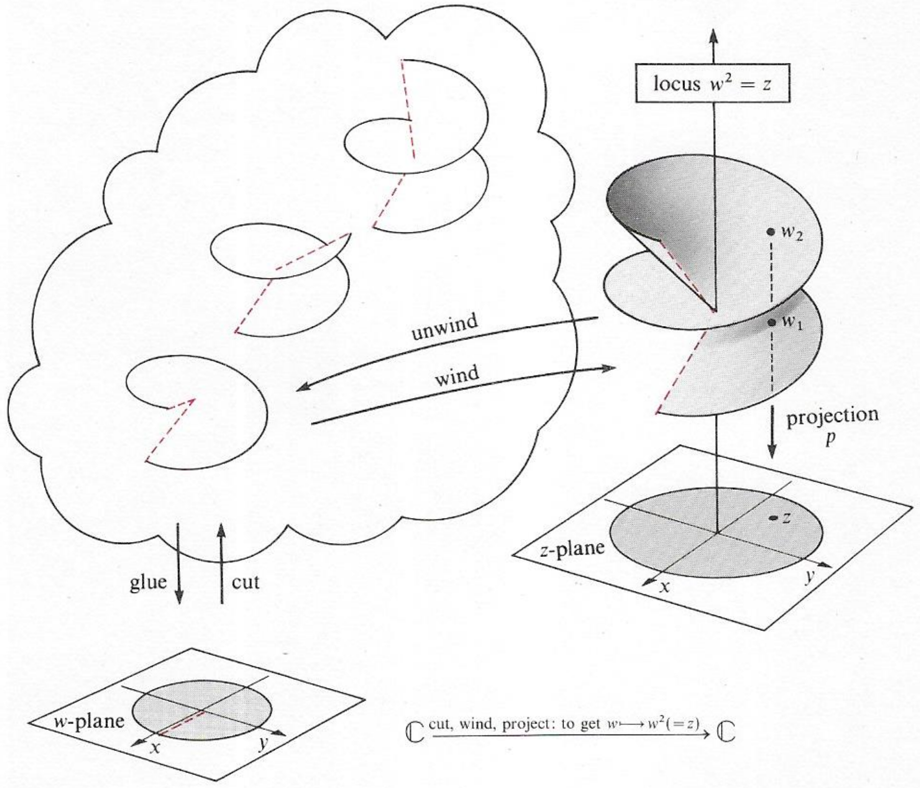
As a prerequisite for watching TV7, you should work through the parts of the audio tape dealing with Frame 1 below. Allow about 20 minutes for this.

The tape section is on side 2 of Cassette A2.

Pre-TV Work

Before watching TV7, revise the idea of "local homeomorphism" in *Unit 6*, then familiarize yourself with Frame 1 before starting the tape.

FRAME 1 **Locus of $w^2 = z$ near $(z, w) = (0, 0)$ and the function $w \longmapsto w^2$**



$$\mathbb{C} \xrightarrow{\text{cut, wind, project: to get } w \longmapsto w^2(=z)} \mathbb{C}$$

(Replacing (z, w) by $\left(\frac{1}{z}, \frac{1}{w}\right)$ converts the locus near (∞, ∞) into this picture.)

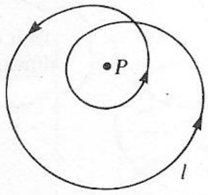
After familiarizing yourself with Frame 1, START THE TAPE.

TV Overview

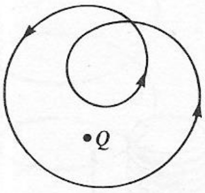
After watching TV7, RESTART THE TAPE.

FRAME 2

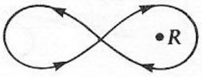
Winding number of a loop about a point



winding number 2



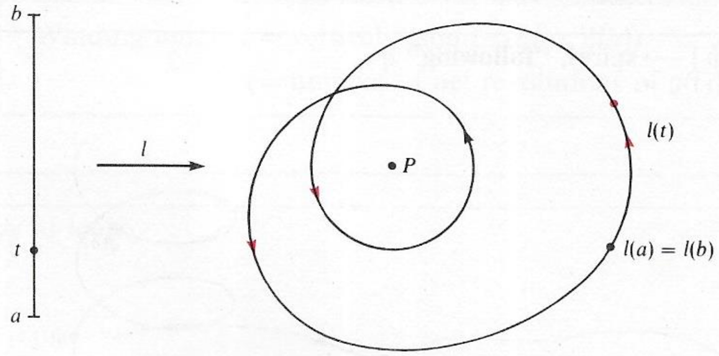
winding number 1



winding number -1

FRAME 3

Definition of loop l in $\mathbb{R}^2$

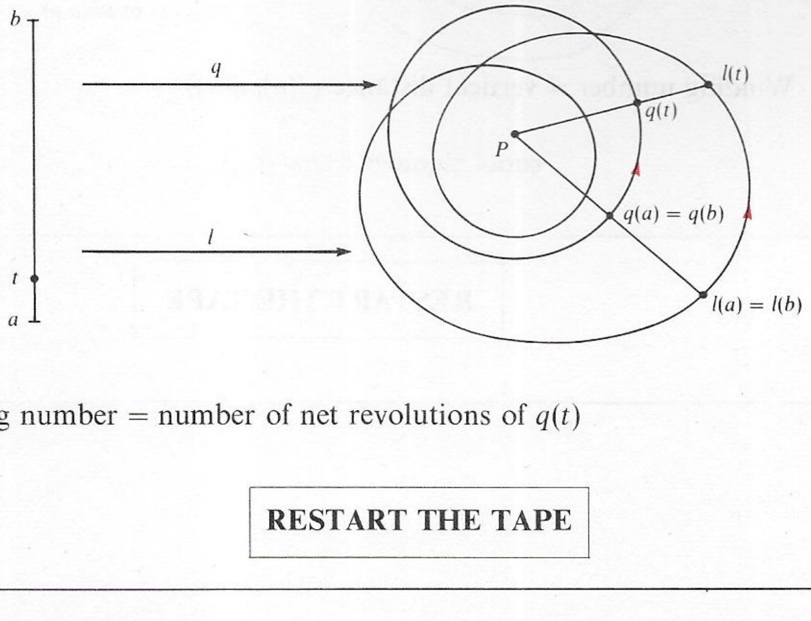

b
t
a
l
P
l(t)
l(a) = l(b)

Continuous $l\colon [a,b] \longrightarrow \mathbb{R}^2$, with $l(a) = l(b)$.

FRAME 4

Continuous $q\colon [a,b] \longrightarrow C$ (circle of directions)

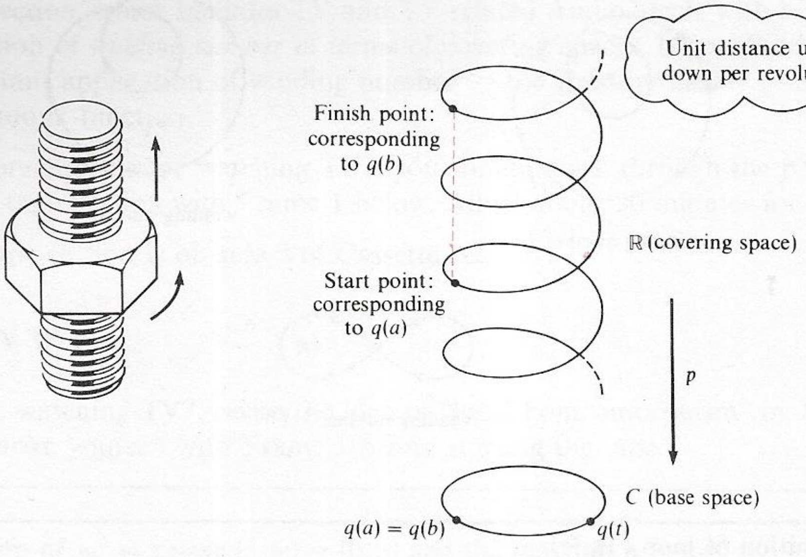
circle C (any radius)


b
q
l
t
a
l(t)
q(t)
P
q(a) = q(b)
l(a) = l(b)

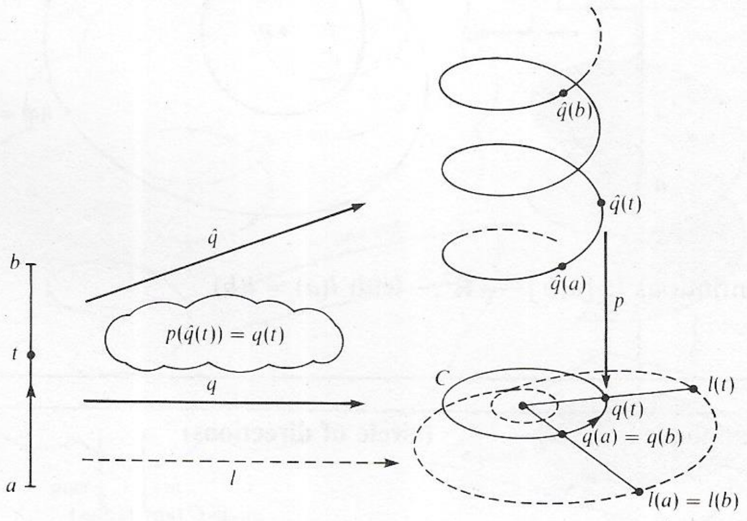
Winding number = number of net revolutions of $q(t)$

RESTART THE TAPE

FRAME 5 **Recording the number of net revolutions**


Unit distance up or down per revolution
Finish point: corresponding to $q(b)$
Start point: corresponding to $q(a)$
$\mathbb{R}$ (covering space)
p
C (base space)
$q(a) = q(b)$
$q(t)$

FRAME 6 $\hat{q}: [a, b] \longrightarrow$ spiral, "following" q

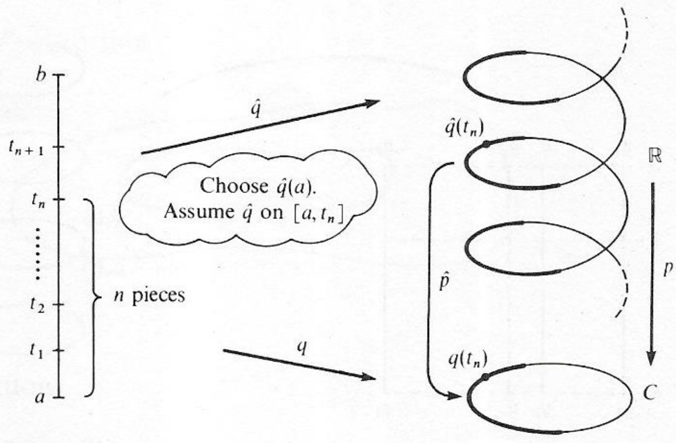

$\hat{q}(b)$
$\hat{q}(t)$
$\hat{q}(a)$
$\hat{q}$
b
t
a
$p(\hat{q}(t)) = q(t)$
q
l
p
C
$q(t)$
$l(t)$
$q(a) = q(b)$
$l(a) = l(b)$

Winding number = vertical distance $(\hat{q}(a), \hat{q}(b))$

RESTART THE TAPE

FRAME 7

Construction of a CONTINUOUS $\hat{q}$

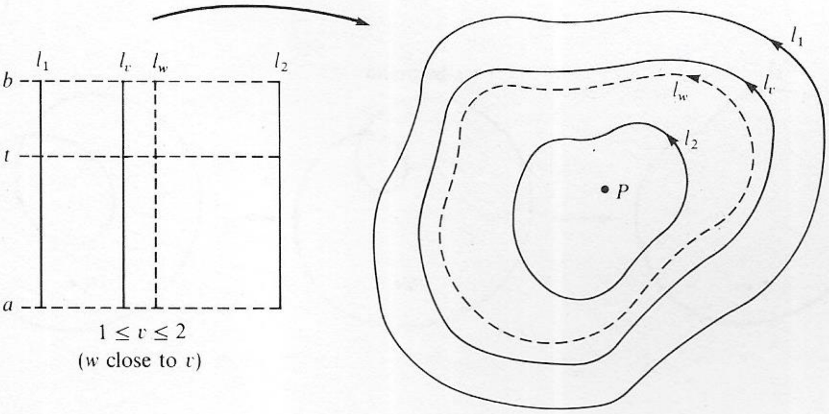

b
t_{n+1}
t_n
t_2
t_1
a
n pieces
$\hat{q}$
Choose $\hat{q}(a)$.
Assume $\hat{q}$ on $[a, t_n]$
q
$\hat{q}(t_n)$
$\hat{p}$
$q(t_n)$
$\mathbb{R}$
p
C

On $[t_n, t_{n+1}]$, $\hat{q}(t) = \hat{p}^{-1}(q(t))$.
By glue rule, $\hat{q}$ continuous on $[a, t_{n+1}]$.

Winding number = vertical distance $(\hat{q}(a), \hat{q}(b))$
(= number of net revolutions of $q(t)$)

FRAME 8

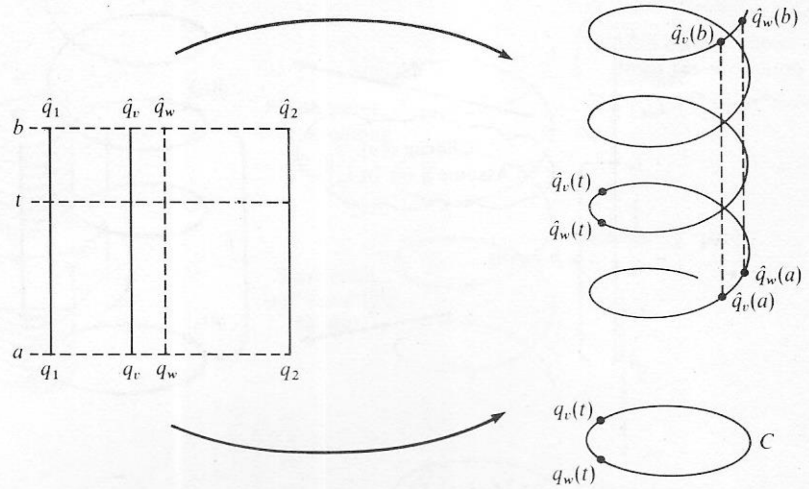
Homotopy of loops


l_1
l_v
l_w
l_2
b
t
a
$1 \leq v \leq 2$
(w close to v)
l_1
l_v
l_w
l_2
P

l_1, l_2 are *homotopic* loops

FRAME 9

Homotopic loops have the same winding number

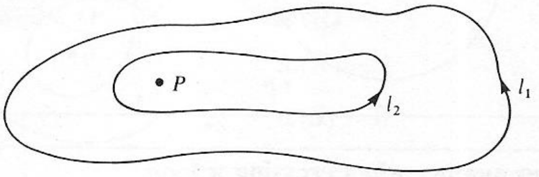

$\hat{q}_1$
$\hat{q}_v$
$\hat{q}_w$
$\hat{q}_2$
b
t
a
q_1
q_v
q_w
q_2
$\hat{q}_v(b)$
$\hat{q}_w(b)$
$\hat{q}_v(t)$
$\hat{q}_w(t)$
$\hat{q}_w(a)$
$\hat{q}_v(a)$
$q_v(t)$
$q_w(t)$
C

Winding number unchanged for small change in v,
i.e. distance $(\hat{q}_v(a), \hat{q}_v(b)) \simeq$ distance $(\hat{q}_w(a), \hat{q}_w(b))$.

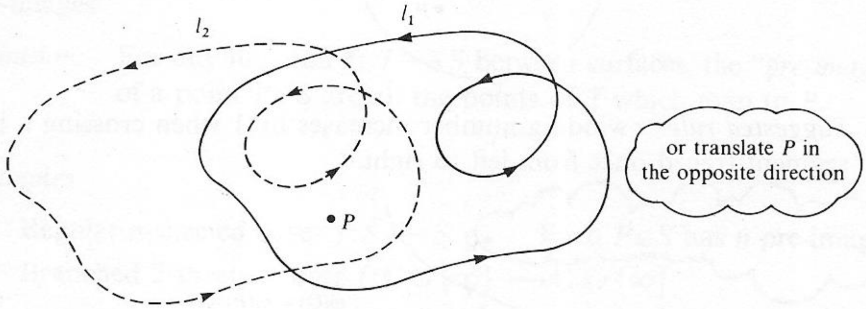
Hence unchanged between $v = 1$ and 2.

FRAME 10 Examples of homotopies

(a) "Elastic" distortion



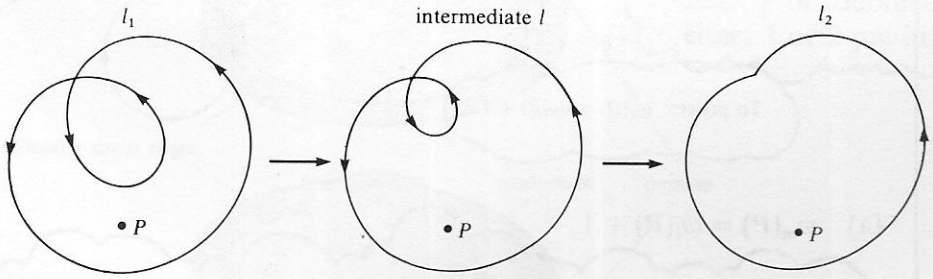
(b) Translation



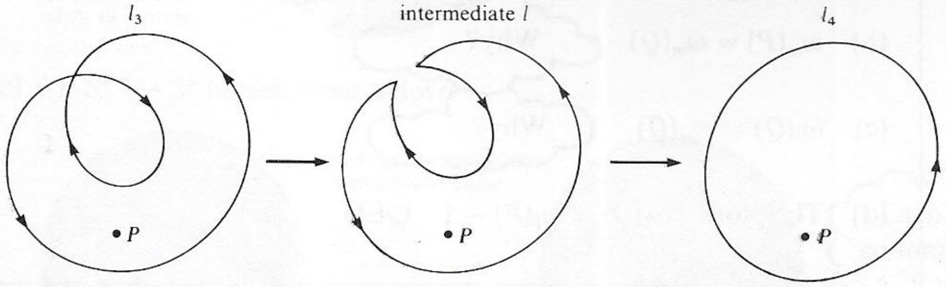
or translate P in the opposite direction

(c) Getting rid of redundant loops

(i)

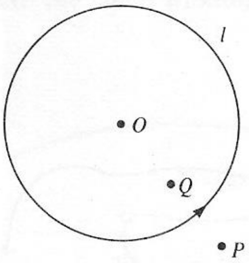

l_1
intermediate l
l_2
P
P
P

(ii)


l_3
intermediate l
l_4
P
P
P

FRAME 11

Winding numbers of a circle (traced once) about different points

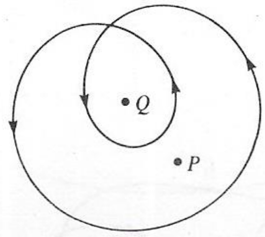


$\omega(O) = \omega(Q) = 1$

$\omega(P) = 0$

FRAME 12

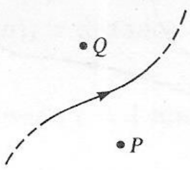
Change in winding number when crossing a loop



$\omega(Q) = 2$

$\omega(P) = 1$

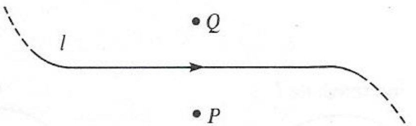
Suggested rule: winding number increases by 1 when crossing a loop segment traced once from left to right.

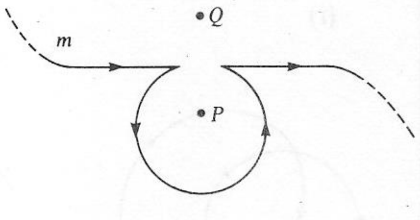


$\omega(Q) = \omega(P) + 1$

FRAME 13

Proof of rule suggested in Frame 12





To prove: $\omega_l(Q) = \omega_l(P) + 1$

(a) $\omega_m(P) = \omega_l(P) + 1$

Imagine constructing a new loop, *n*, by deforming *m* until its "horns" touch. Then $\omega_m(P) = \omega_n(P)$ by homotopy. But *n* circles *P* once more than does *l*. QED

(b) $\omega_m(P) = \omega_m(Q)$ Why?

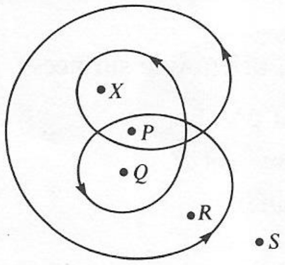
(c) $\omega_l(Q) = \omega_m(Q)$ Why?

(d) Therefore $\omega_l(Q) = \omega_l(P) + 1$ QED

After working through the above, and completing parts (b) and (c),
RESTART THE TAPE

FRAME 14

“Rule of thumb” calculation of winding numbers



$\omega(X) =$

$\omega(P) =$

$\omega(Q) =$

$\omega(R) = 1$

$\omega(S) = 0$

After filling in the blanks,
RESTART THE TAPE

FRAME 15

Pre-images

Definition: For any function $f\colon T \longrightarrow S$ between surfaces, the “*pre-images*” of a point $P \in S$ are all the points of T which map to P.

Examples

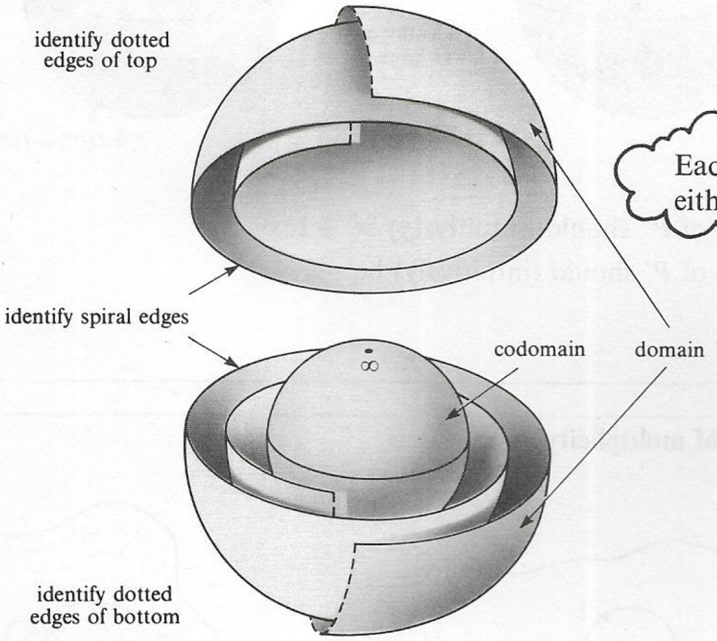
(a) Regular n-sheeted cover $f\colon \tilde{S} \xrightarrow{n} S$

Each $P \in S$ has n pre-images

(b) Branched 2-sheeted cover $f\colon \mathbb{C} \cup \{\infty\} \longrightarrow \mathbb{C} \cup \{\infty\}$

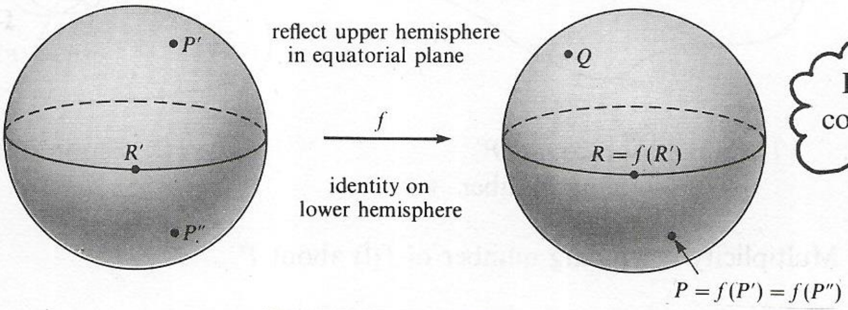
$$w \longmapsto w^2$$

(as in Pre-TV Work, Frame 1)


identify dotted edges of top
identify spiral edges
codomain
domain
∞
identify dotted edges of bottom

Each P of codomain has either 1 or 2 pre-images

(c) $f\colon S^2 \longrightarrow S^2$ as described below


P′
R′
P″
reflect upper hemisphere in equatorial plane
f
identity on lower hemisphere
Q
R = f(R′)
P = f(P′) = f(P″)

Each point of the codomain has 0, 1 or 2 pre-images

RESTART THE TAPE

FRAME 16

Brouwer Degree Theorem

Continuous $f\colon S_1 \longrightarrow S_2$ (S_1, S_2 closed, orientable surfaces)

For each P in S_2 with a finite number of pre-images,

N_P = sum of multiplicities of all pre-images of P.

(The term multiplicities will be defined later.)

Theorem

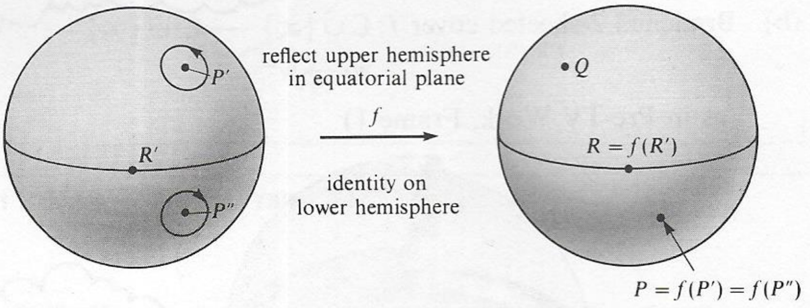
If every P in S_2 has a finite number of pre-images, N_P is the same for all P.

(This value N_P is called the Brouwer degree of f.)

RESTART THE TAPE

FRAME 17

Multiplicities in Example (c) of Frame 15



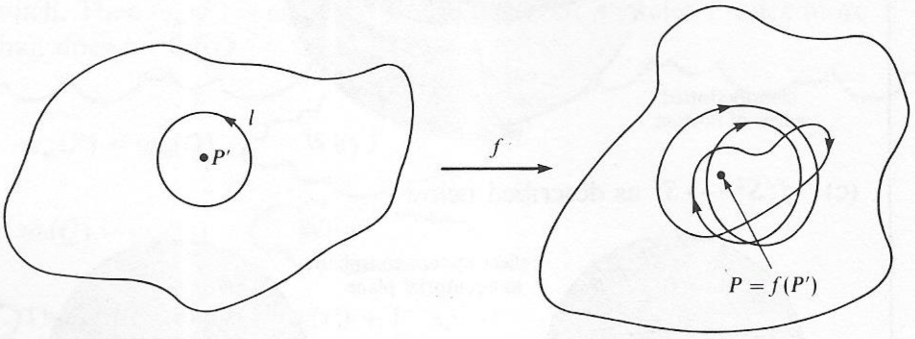
P'
R'
P''
reflect upper hemisphere in equatorial plane
f
identity on lower hemisphere
Q
$R = f(R')$
$P = f(P') = f(P'')$

Multiplicity of P'' should (intuitively) be $+1$.

Multiplicity of P' should (intuitively) be -1.

FRAME 18

Definitions of multiplicity



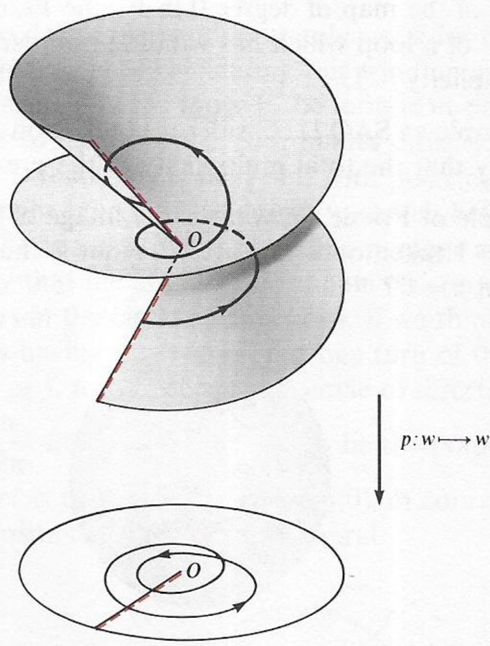
l = small loop round P' with winding number $+1$

$f(l)$ = loop round P

Multiplicity = winding number of $f(l)$ about P

FRAME 19

Multiplicity of a branch point for Example (b) of Frame 15


O
$p: w \longmapsto w^2$
O

Multiplicity of O is $+2$.

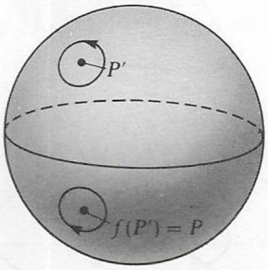
Post-TV Work

SAQ 1 For the example of the map of degree 0 in Frame 17, consider the point P'. Sketch the image of a loop which has winding number $+1$ about P', and check that P' has multiplicity -1.

SAQ 2 In the same example as SAQ 1, consider a point R on the equator. What is its pre-image? Verify that the total multiplicity of the pre-image points is zero.

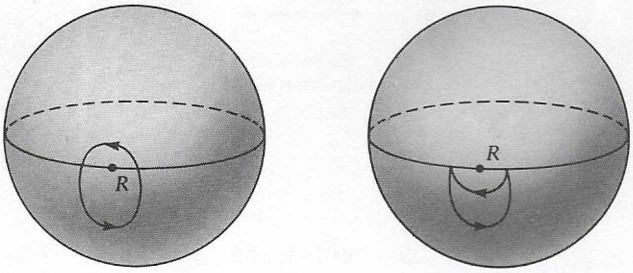
SAQ 3 In the last example of Frame 19, what is the image of the loop given by the boundary of $|z| = 1$ taken once positively? What is the winding number of this image loop about $z = 0$?

SOLUTION TO SAQ 1



The orientation is reversed by reflection, so the image loop goes round in the opposite sense. The winding number is therefore -1 if the original loop has winding number $+1$. So the multiplicity of P' is -1. ■

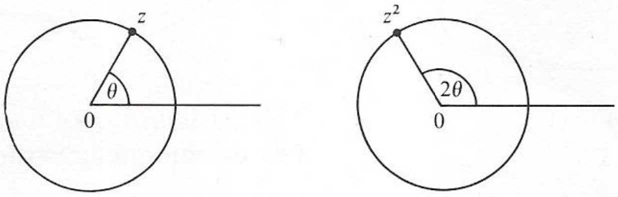
SOLUTION TO SAQ 2



The pre-image of R is just R itself. To verify that the multiplicity of R is zero, take a small loop round R. Draw an unsymmetrical loop (not a circle centre R) for convenience first. The image loop is shown on the right and has winding number 0 about R = image of R.

If we had taken a circle centre R, the image loop would have been the semicircle taken twice, once in each direction. ■

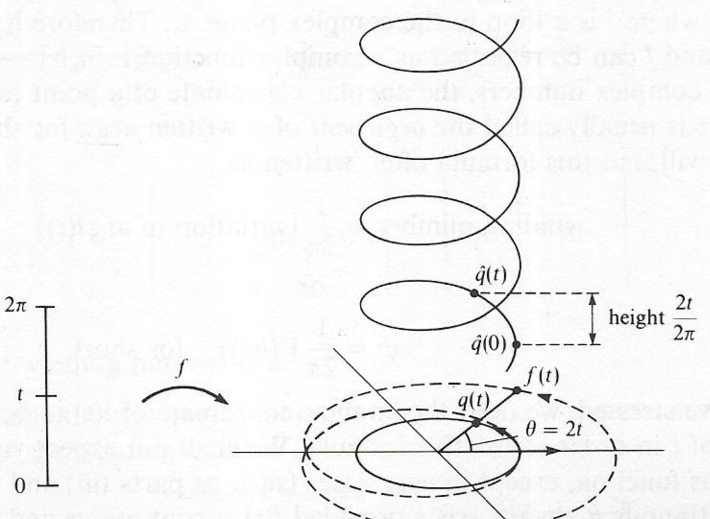
SOLUTION TO SAQ 3 The image is the loop given by the circle $|z| = 1$ taken twice positively; i.e. as a point runs round the circle $|z| = 1$ *once*, the image point runs round the circle $|z| = 1$ twice in the same direction. Algebraically, you can see that the point $z = e^{it}$ is sent by $z \longmapsto z^2$ to the point $z^2 = e^{2it}$ so, as the ray joining 0 to z rotates, the ray joining 0 to the image point z^2 rotates twice as fast. The winding number of the image loop about $z = 0$ is 2.



■

7.2 A Formula for the Winding Number

We have seen how to define the winding number of a loop about a point, and have developed a simple way of calculating this winding number. But this method depends on sketching the loop. If the loop is in polar coordinate form, it is sometimes easier to use the polar coordinates directly. For instance, the loop $f\colon [0, 2\pi] \longrightarrow \mathbb{R}^2$, where $f(t)$ is the point with polar coordinates $(2 + t(2\pi - t), 2t)$, might take us a little time to sketch before we could use the methods of Section 7.1 to work out the winding number about the origin. But we can see straight away that the angular coordinate $\theta(t)$ (in our example $\theta(t) = 2t$) gives us the point $q(t)$ in the circle of directions. If we think of the spiral above the circle of directions as having pitch 1, so that one turn of the spiral corresponds to a vertical movement of 1, a turn of α in the circle of directions corresponds to a change in height of $\dfrac{\alpha}{2\pi}$. So we can take $\hat{q}(t)$ to be the point which is $\theta(t)/2\pi$ above $\hat{q}(0)$. This point is not *vertically* above $\hat{q}(0)$, of course, but its height above $\hat{q}(0)$ determines its position uniquely on the spiral.



So our formula for the winding number can be written directly in terms of the angular coordinate $\theta(t)$, as follows. For a loop $f\colon [a, b] \longrightarrow \mathbb{R}^2$,

$$\begin{aligned}\text{winding number} &= \text{height of } \hat{q}(b) - \text{height of } \hat{q}(a)\\ &= \theta(b)/2\pi - \theta(a)/2\pi\\ &= \frac{1}{2\pi}(\theta(b) - \theta(a)).\end{aligned}$$

SAQ 4 Find the winding numbers about 0 of the following loops.

(i) $l(t) = (2 + \sin t, -t)$ (polar coordinates, $0 \le t \le 2\pi$

(ii) $l(t) = (\frac{3}{2} + \cos t, 2t + \pi)$ (polar coordinates), $0 \le t \le 2\pi$

(iii) $l(t) = (\cos 3t, \sin 3t)$ (Cartesian coordinates), $0 \le t \le 2\pi$

(iv) $l(t) = (\cos 4\pi t, \sin 4\pi t)$ (Cartesian coordinates), $0 \le t \le 1$

SOLUTION TO SAQ 4 (i) We can take $\hat{q}(t) =$ point of height $\dfrac{-t}{2\pi}$ above $\hat{q}(0)$, and so the change in height of $\hat{q}(2\pi)$ is -1, which is the winding number. More simply, $\theta(t) = -t$, and

$$\text{winding number} = \frac{1}{2\pi}(\theta(2\pi) - \theta(0)) = \frac{1}{2\pi}(-2\pi - 0) = -1.$$

(ii) Here $\theta(t) = 2t + \pi$, so

$$\text{winding number} = \frac{1}{2\pi}(2 \times 2\pi + \pi - (2 \times 0 + \pi)) = 2.$$

(iii) In polar form, we can write this as $l(t) = (1, 3t)$, so

$$\text{winding number} = \frac{1}{2\pi}(3 \times 2\pi - 0) = 3.$$

(iv) Because t varies from 0 to 1, the angular coordinate of $l(t)$ varies from 0 to 4π, and the winding number is 2. ■

The expression $\theta(b) - \theta(a)$ which occurs in the winding number formula is called the *variation in the angular coordinate* of l. In **B**, you will find several examples where l is a loop in the complex plane, $\mathbb{C}$. Therefore $l(t)$ is a complex number, and l can be regarded as a complex function $l\colon [a,b] \longrightarrow \mathbb{C}$. In the theory of complex numbers, the angular coordinate of a point (complex number) z is usually called the *argument* of z, written $\arg z$ for short. Therefore in **B** you will find this formula often written as

$$\text{winding number} = \frac{1}{2\pi}(\text{variation in } \arg l(t))$$

or

$$\omega = \frac{1}{2\pi}V(l(t)), \quad \text{for short.}$$

As we have stressed, we need the angular coordinate of $l(t)$ to be a *continuous* function of t in order to use this formula. We shall not expect you to *find* a continuous function, except in easy cases (such as parts (iii) and (iv) of SAQ 4), but a continuous θ always exists provided $l(t)$ is continuous and does not pass through 0. (A proof of this is given in the (optional) Appendix II, page 31.)

The following example is slightly different. It shows how we can sometimes avoid sketching the loop and still avoid finding an explicit formula for the angular coordinate $\theta(t)$.

READ B page 114, Example 1, to page 115, line 5.

NOTE *Page 115, line 3*
(Solution of $x = 0$) If $x = \cos t + \cos 2t - 1 = 0$ and $0 \leq t \leq 2\pi$, we can use the identity $\cos 2t = 2\cos^2 t - 1$ to obtain $2\cos^2 t + \cos t - 2 = 0$. Solving this quadratic gives $\cos t = \frac{1}{4}(-1 \pm \sqrt{17})$. Since $\frac{1}{4}(-1 - \sqrt{17}) < -1$, we need only look at the solutions of $\cos t = \frac{1}{4}(-1 + \sqrt{17})$. This has two solutions, one in the first and one in the fourth quadrant. ▲

SAQ 5 **B** page 130, Question 2.

SOLUTION TO SAQ 5 If we take the translation of the plane taking $(-\frac{1}{2}, 0)$ to the origin, then the winding number about $(-\frac{1}{2}, 0)$ of the given curve is the same as the winding number about the origin of the curve given by

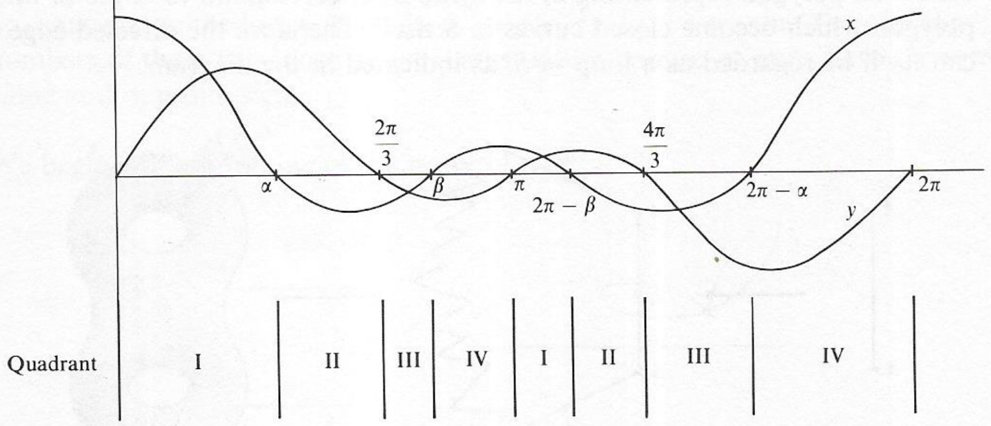
$$\left.\begin{aligned} x &= \cos t + 2t + \tfrac{1}{2}, \\ y &= \sin t + \sin 2t, \end{aligned}\right\} \quad 0 \le t \le 2\pi.$$

As in **B**'s Example 1, $y = 0$ at $t = 0$, $2\pi/3$, π, $4\pi/3$ and 2π. The zeros of x are given by $\cos t + \cos 2t + \frac{1}{2} = 0$, i.e. by $2\cos^2 t + \cos t - \frac{1}{2} = 0$, so

$$\cos t = \frac{-1 \pm \sqrt{5}}{4}.$$

So x has four zeros in $[0, 2\pi]$, at α, β, $2\pi - \beta$ and $2\pi - \alpha$, where

$$\alpha = \arccos\left(\frac{-1+\sqrt{5}}{4}\right), \ \beta = \arccos\left(\frac{-1-\sqrt{5}}{4}\right) \text{ or } \alpha \sim 1.26, \beta \sim 2.51.$$



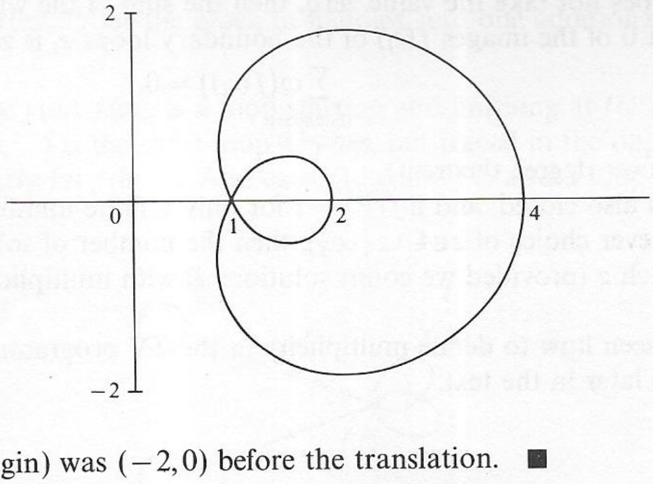
Hence the winding number is 2.

To determine the winding number about $(-2, 0)$ of the same curve, we investigate the winding number about the origin of

$$\left.\begin{aligned} x &= \cos t + \cos 2t + 2, \\ y &= \sin t + \sin 2t, \end{aligned}\right\} \quad 0 \le t \le 2\pi.$$

For any t, $\cos t \ge -1$, and therefore $x \ge -1 - 1 + 2 = 0$. This means that the curve is confined to the 1st and 4th quadrants, and so the winding number is 0.

You might like to sketch the curve and thus calculate these winding numbers by inspection. However, the sketching is not altogether straightforward, whereas the method suggested has the advantage of being reasonably systematic for any "parametric" curve. For interest, the curve looks like this.



0 (the new origin) was $(-2, 0)$ before the translation. ■

7.3 Proof of the Brouwer Degree Theorem

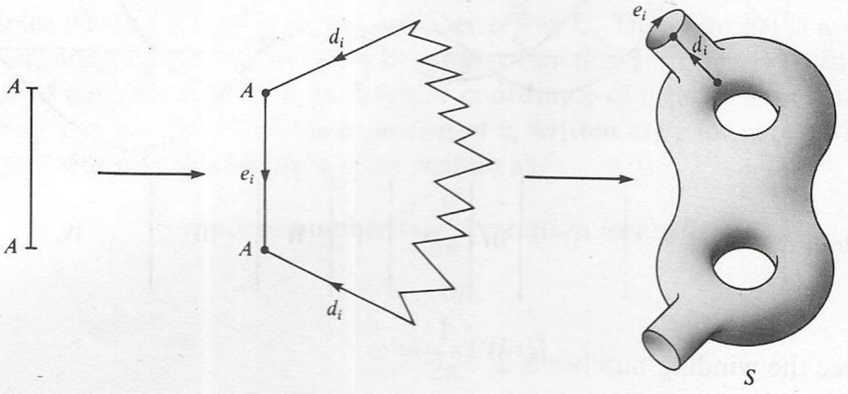
The Brouwer degree theorem uses properties of winding number to prove results about functions. The functions we are going to consider will have as domain a surface S, and as codomain the Riemann sphere $\mathbb{C} \cup \{\infty\}$. As usual, $(x, y) \in \mathbb{R}^2$ corresponds to $z = x + iy \in \mathbb{C}$.

A loop on S is a continuous function $l: [a, b] \longrightarrow S$, where $[a, b]$ is an interval and $l(a) = l(b)$. If $f: S \longrightarrow \mathbb{C} \cup \{\infty\}$ is any continuous function, the composite function $f \circ l$ (or $f(l)$ as **B** writes it) is a loop on the Riemann sphere. Hence we can talk about winding numbers of $f(l)$ about points on the Riemann sphere.

To start with, let us look at an orientable surface S whose canonical equation is

$$a_1 b_1 a_1^{-1} b_1^{-1} \dots a_p b_p a_p^{-1} b_p^{-1} d_1 e_1 d_1^{-1} \dots d_r e_r d_r^{-1} = 1.$$

S has p handles and r boundary curves, $e_1, \dots, e_r$. In fact, if we take the canonical polygon representing S, the symbols e_i correspond to edges of this polygon which become closed curves in S itself. Therefore the directed edge e_i can itself be regarded as a loop in S, as indicated in the diagram.



representing polygon

For any function $f: S \longrightarrow \mathbb{C}$ which never takes the value zero, we can therefore consider the image of this loop, $f(e_i)$, and calculate its winding number about zero. We shall see that the sum of these winding numbers over all boundary loops $e_1, \dots, e_r$ has to be zero. More precisely, we shall show that the following is true.

THEOREM Let $f: S \longrightarrow \mathbb{C} \cup \{\infty\}$ be a continuous function from an orientable surface S to the Riemann sphere

(i) If f does not take the value zero, then the sum of the winding numbers about 0 of the images $f(e_i)$ of the boundary loops e_i is zero; that is

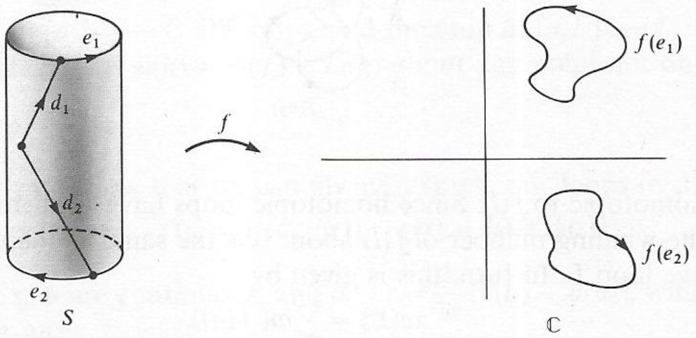
$$\sum_{\substack{e_i \text{ (boundary} \\ \text{curve of } S)}} \omega(f(e_i)) = 0.$$

(ii) (Brouwer degree theorem)
If S is also closed, and if $f(P) = z$ for only a finite number of points P of S (whatever choice of $z \in \mathbb{C} \cup \{\infty\}$), then the number of solutions is the same for each z (provided we count solutions P with multiplicities).

You have seen how to define multiplicity in the TV programme; the definition is repeated later in the text.

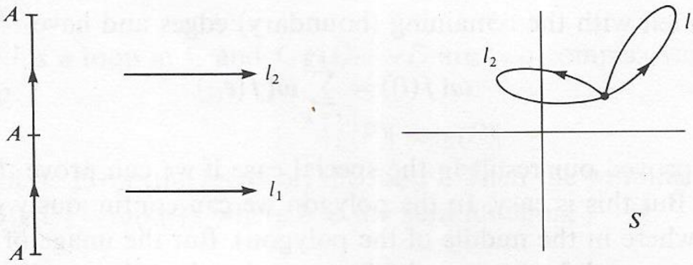
EXAMPLE 1 (Illustrating part (i) of the Theorem)

Suppose S is the cylinder whose canonical edge equation is $d_1e_1d_1^{-1}d_2e_2d_2^{-1} = 1$. Any map $f: S \longrightarrow \mathbb{C}$ gives two loops, $f(e_1)$ and $f(e_2)$, in $\mathbb{C}$. If $f(S)$ does not include 0, then since $f(e_1)$ and $f(e_2^{-1})$ are homotopic, they have the same winding number about 0.



So, considering the loops $f(e_1)$, $f(e_2)$ $[= f(e_2^{-1})^{-1}]$, we see that the winding numbers of these two loops about 0 must either be zero or must have the same value and opposite sign. □

We begin with an outline of the proof of (i).



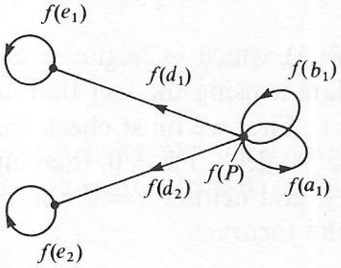
First, we need to notice that given two loops $l_1: [a, b] \longrightarrow S$ and $l_2: [b, c] \longrightarrow S$, we can form a new loop $l: [a, c] \longrightarrow S$ by glueing the two loops together, provided that the initial point, $l_2(b)$, of l_2 is the same as the final point, $l_1(b)$, of l_1. The new function, l, will then be continuous, by the glue rule. And it is obvious that $\omega(l) = \omega(l_1) + \omega(l_2)$, where we use $\omega(l)$ to denote the winding number of l about some point not on l.

Now look at the polygon representing S in canonical form. The boundary enables us to define a loop in S, l say, and this loop is mapped by f to $\mathbb{C} \cup \{\infty\}$. In fact, $f(l)$ consists of several other loops starting and finishing at the same point. To see this, notice that the canonical equation for S is

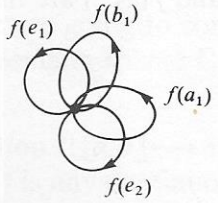
$$a_1b_1a_1^{-1}b_1^{-1} \ldots a_pb_pa_p^{-1}b_p^{-1}d_1e_1d_1^{-1} \ldots d_re_rd_r^{-1} = 1.$$

This has one vertex, P, for the block of handles, and one additional vertex for each cuff.

Now, each of the $f(a_i)$, $f(b_i)$, is a loop starting and finishing at $f(P)$. And, furthermore, $f(a_i^{-1})$ is the same loop as $f(a_i)$, but traced in the opposite direction. Similarly for $f(b_i^{-1})$. And each $f(d_je_jd_j^{-1})$ is also a loop based at $f(P)$.



Since f never takes the value 0, by a homotopy we can shrink the "stalks" $f(d_i)$ to obtain the loop L,



which is homotopic to $f(l)$. Since homotopic loops have the same winding number, the winding number of $f(l)$ about 0 is the same as the winding number of the above loop L. In turn, this is given by

$$\omega(L) = \sum_{\substack{\text{all edges } g \\ \text{except } d_i\text{s}}} \omega(f(g)).$$

Now since, for example,

$$\omega(f(a_i)) = -\omega(f(a_i^{-1})),$$

the effect of all the a_is and b_is in the above sum is to cancel each other and give zero.

Thus we are left with the remaining (boundary) edges and have

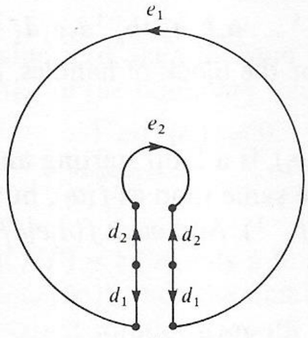
$$\omega(f(l)) = \sum_{i=1}^{r} \omega(f(e_i)).$$

So we have proved our result in the special case if we can prove that $\omega(f(l)) = 0$. But this is easy. In the polygon we can continuously shrink l into a point (somewhere in the middle of the polygon). But the image of l as it shrinks gives a continuous deformation of $f(l)$ to a point. And the winding number of the "point loop" is zero, about any point! Of course, we have to use the fact that f does not map any point of S to zero — otherwise we could not be sure that the loop $f(l)$ did not pass through 0 as it shrank, even though 0 was not on the original loop $f(l)$.

EXAMPLE 2 **B** page 130, Section 4.2, Question 1.

Solution

The canonical form of the edge equation of the annulus is $d_1 e_1 d_1^{-1} d_2 e_2 d_2^{-1} = 1$, corresponding to the figure below.



We are asked for $V(f(e_1))$, which is 2π times the winding number of $f(e_1)$ about 0. We choose to calculate it using the fact that $\omega(f(e_1)) + \omega(f(e_2)) = 0$, so $V(f(e_1)) + V(f(e_2)) = 0$. First we must check that f does not vanish on the annulus. But if $f(z) = z^2 + (|z| - 1)z = 0$, then either $z = 0$ or $|z| + z - 1 = 0$. The latter implies $z = \frac{1}{2}$, and neither $z = 0$ nor $z = \frac{1}{2}$ lies in the annulus. Therefore we can use the theorem.

On e_2, we have $|z| = 1$, so $f(z) = z^2$. Since e_2 is described once *clockwise*, $\omega(f(e_1)) = -2$. By the theorem, $\omega(f(e_1)) = 2$, i.e. $V(f(e_1)) = 4\pi$. □

To complete the proof of Brouwer's theorem we need a couple of results about winding numbers of loops in $\mathbb{C}$.

LEMMA Suppose $f, g\colon [a,b] \longrightarrow \mathbb{C}$ are loops, and the loop $h\colon [a,b] \longrightarrow \mathbb{C}$ is defined by $h(t) = f(t)g(t)$. Then $\omega(h) = \omega(f) + \omega(g)$ about any point not on the loops.

Proof

On page 31, we show that we can always express our loops in the polar form

$$f(t) = (r(t), \theta(t)), \quad g(t) = (s(t), \phi(t)),$$

where r, θ, s, ϕ are continuous, and $\omega(f) = \dfrac{1}{2\pi}(\theta(b) - \theta(a))$, with a similar formula for $\omega(g)$.

By the formula for the product of two complex numbers in polar form, we have

$$h(t) = (r(t)s(t), \theta(t) + \phi(t)),$$

so

$$\omega(h) = \frac{1}{2\pi}(\theta(b) + \phi(b) - \theta(a) - \phi(a)) = \omega(f) + \omega(g), \quad \text{as required.} \quad \blacksquare$$

THEOREM **(Rouché)**

Suppose l is a loop in $\mathbb{C}$ and $f, g\colon \mathbb{C} \longrightarrow \mathbb{C}$ are two complex-valued functions satisfying

$$|f(P)| < |g(P)|$$

for all points P on (the locus of) the loop l. Then the winding numbers about 0 satisfy $\omega(g(l)) = \omega(h(l))$, where h is the sum function, $f + g$.

Proof

Because $|f(P)| < |g(P)|$, we cannot have $g(P) = 0$ for any point P on l.

Consider

$$\begin{aligned} h(P) = f(P) + g(P) &= g(P)[1 + f(P)/g(P)] \\ &= g(P)\bar{f}(P), \quad \text{say,} \end{aligned}$$

if we write $\bar{f}(P)$ for the expression in square brackets.

Take winding numbers about 0. Our lemma tells us that

$$\omega(h) = \omega(g) + \omega(\bar{f}).$$

But $|f(P)/g(P)| < 1$, so $\bar{f}(P) = 1 + f(P)/g(P)$ stays *inside* the circle $|z - 1| = 1$. So $\omega(\bar{f}) = 0$. ■

EXAMPLE 3 Calculate $\omega(h(l))$, where l is the loop consisting of the circle $|z| = 2$ taken once, and $h(z) = z^2 + z + 1$.

Solution

On l, $4 = |z^2| > 3 = |z| + 1 \geq |z + 1|$. Let $f(z) = z + 1$ and $g(z) = z^2$. Then $|f(z)| < |g(z)|$ on l, and we can apply Rouché's theorem to deduce that $\omega(g(l)) = \omega(h(l))$. We know that $\omega(g(l))$ is 2, so $\omega(h(l))$ is also 2. ■

SAQ 6 l is the loop given in Example 3, and $h(z) = z^3 - z^2 + z - 1$. Find $\omega(h(l))$ about 0.

SAQ 7 By applying Rouché's theorem, calculate $\omega(f(l_1))$, $\omega(f(l_2))$ about 0, where $f(z) = z^4 + 2z + 4$, l_1 is the unit circle $|z| = 1$, and l_2 is the circle $|z| = 2$, both taken once positively.

SOLUTION TO SAQ 6 On l, we have

$$8 = |z^3| > |z^2| + |z| + 1 \geq |-z^2 + z - 1|;$$

so $\omega(h(l)) = \omega(g(l))$, where $g(z) = z^3$,
$= 3.$ ■

SOLUTION TO SAQ 7 (i) If $|z| = 1$, then $|2z| = 2$, and so $|2z + 4|$ cannot be less than 2. (As z varies on $|z| = 1$, $2z + 4$ traces out a circle centre 4, radius 2, and the nearest point of this circle to the origin is 2.) Therefore $|z^4| < |2z + 4|$ on the boundary circle $|z| = 1$, and so $f(l_1)$ has the same winding number about 0 as has $g(l_1)$, where $g(z) = 2z + 4$. This has winding number 0.

(ii) If $|z| = 2$, then $|z^4| = 16$ and $|2z + 4| \leq 8$. So this time $f(l_2)$ has the same winding number about 0 as has $h(l_2)$, where $h(z) = z^4$. This has winding number 4. ■

There are more examples of this nature in the Exercise Booklet.

We can now turn to part (ii) of our theorem, usually called the Brouwer degree theorem. (The proof is *not assessable*, but it is worth reading as it is ingenious.) The idea of the proof is to cut out a disc from S around each point P that maps to a given point — say 0—of $\mathbb{C} \cup \{\infty\}$. Then on this *punctured* surface with discs removed, the function f never takes the value 0. We now apply part (i) of the theorem to this new surface. The boundary consists of (a) the boundary of S, together with (b) the boundaries of all the discs which have been removed. So part (i) of the theorem allows us to relate the winding numbers of the images of the disc boundaries (b) to what happens on the boundary of S.

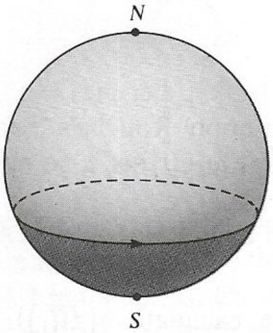
READ B Section 4.3 beginning on page 120.

NOTES (i) *Page 120, lines 21–2*
By a *disc* in S we mean a homeomorphic copy of a disc in $\mathbb{R}^2$.

(ii) *Page 120, line -12*
B says that the edges in the canonical equation can be chosen to miss the disc. We suggest that you assume this is the case — it is not implausible that one can just move the edges of the polygon a small amount to avoid the discs.

(iii) *Page 120, line -8; Page 121, line -5*
B has made a slight mistake here. In fact, the canonical form of an orientable surface given on page 120, line -13, contains $r + 1$ vertices—one for the block of handles and one for each cuff. Here, **B** is referring to the vertex corresponding to all the handles.

(iv) *Page 122, line −3 to page 123, line 1*



A disc on the sphere with boundary (say) a parallel of latitude containing the south pole, S, is shown shaded. The disc boundary is oriented clockwise, as it would appear to an observer at S. But the same parallel of latitude is the boundary of the unshaded disc, which contains the north pole, N, and as a disc boundary of the unshaded disc, this parallel of latitude has the opposite orientation, i.e. as it appears to an observer at N.

(v) *Page 123, lines −7 to −3*

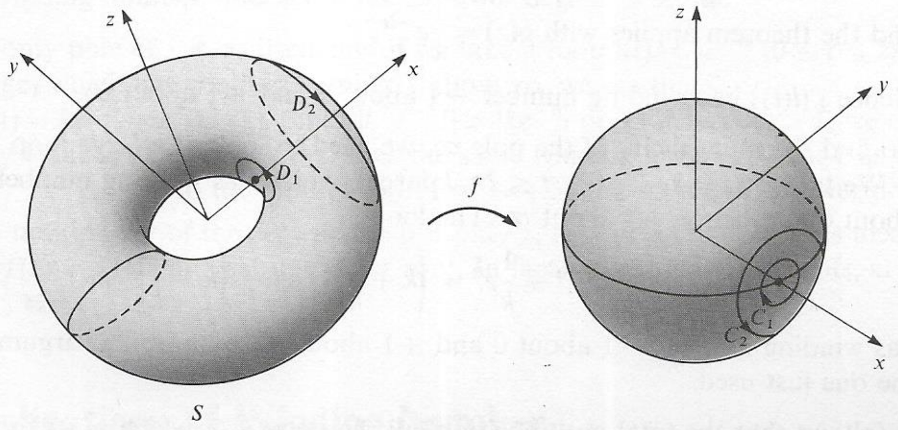
Since f is continuous, for any point P close to a pole Q of f, we have $f(P)$ near ∞, and so $\dfrac{1}{|f(P)|}$ is near 0. This means $|f(P)|$ is large; by taking P near enough to Q we can ensure $|f(P)| > |z_0|$. Choose C to be a small loop around Q so that $|f(P)| > |z_0|$ for every P on C.

By an application of Rouché's theorem, the winding numbers of $f(l)$ and $(f - z_0)(l)$ are the same, which shows that the multiplicities of the pole are the same for f and for h. ▲

EXAMPLE 4 **B** page 131, Question 3(a) on Section 4.3.

Solution

$Q = (1, 0, 0)$, and so if $f(P) = (x/r, y/r, z/r) = Q$, we have $x = r$ and $y = z = 0$. Therefore the only points mapped to Q are $(+\lambda, 0, 0)$ with $\lambda > 0$. Now $(\lambda, 0, 0)$ lies on S for $\lambda = \pm 1$ or $\lambda = \pm 5$; so the points mapped to Q are $(1,0,0)$ and $(5,0,0)$.



We take small discs D_1 around $(1,0,0)$ and D_2 around $(5,0,0)$, and orient their boundaries positively: the positive direction is the anticlockwise direction as seen from outside the torus S from just above the discs. (Arrows in the figure show the boundary direction.) The images under f of the boundaries are C_1 and C_2, with the directions as given. We can see that C_1 is taken negatively on the sphere, C_2 positively, and each winds once around $Q = (1,0,0)$. Therefore

$$\mu[(1, 0, 0), Q] = -1, \quad \mu[(5, 0, 0), Q] = +1.$$

It is easy to show that f maps no point of S to $(0,0,1)$. This can be done algebraically as above, or geometrically: f maps each point P of S to the point on the sphere, S_0, where the halfline joining P to O meets S_0. There are no points of S on the z-axis, so $(0,0,1)$ is not in the image.

The Brouwer degree is 0, as can be seen from adding the multiplicities of the points which map to Q or from the fact that no points are mapped to R.

EXAMPLE 5 Find the zeros and poles of the function $f\colon \mathbb{C} \cup \{\infty\} \longrightarrow \mathbb{C} \cup \{\infty\}$ given by

$$f(z) = \begin{cases} z + \dfrac{1}{z}, & \text{if } z \neq 0, \infty, \\ \infty, & \text{if } z = 0 \text{ or } \infty. \end{cases}$$

Find also the multiplicities of the poles, and hence the Brouwer degree. What is the multiplicity of the point 2?

Solution

Zeros If $z + \frac{1}{z} = 0$, then $z^2 + 1 = 0$ and so $z = \pm i$.

Poles If $z + \frac{1}{z} = \infty$, then either $z = 0$ or $z = \infty$ (by definition of f).

Multiplicity of the pole at $z = 0$ We have to take a small loop l winding once round $z = 0$, and calculate the winding number of $f(l)$ about ∞. This is the same as *minus* the winding number of $f(l)$ about 0. So we take

$$l(t) = ke^{it}, \qquad 0 \leq t \leq 2\pi \quad (k \text{ small}),$$

and look at the winding number of $f(l)$ about 0.

$$\begin{aligned} f\big(l(t)\big) &= ke^{it} + \frac{1}{k}e^{-it} \\ &= \left(\frac{1}{k} + k\right)\cos t - \left(\frac{1}{k} - k\right) i \sin t. \end{aligned}$$

This has the same winding number about 0 as does $\frac{1}{k}(\cos t - i \sin t) = \frac{1}{k}e^{-it}$. We can see that this is true by Rouché's theorem: if k is small,

$$\left|\frac{1}{k}e^{-it}\right| > |ke^{it}|$$

and the theorem applies with $g(z) = \frac{1}{k}e^{-iz}$.

Hence $f\big(l(t)\big)$ has winding number -1 about 0 and $+1$ about ∞.

To find the multiplicity of the pole ∞, we need to look at a *large* loop l round 0. We take $f(t) = ke^{-it}$ $(0 \leq t \leq 2\pi$, k large), which has winding number -1 about 0 and hence $+1$ about ∞. The loop

$$f\big(l(t)\big) = ke^{-it} + \frac{1}{k}e^{it} = \left(k + \frac{1}{k}\right)\cos t - \left(k - \frac{1}{k}\right) i \sin t$$

has winding number -1 about 0 and $+1$ about ∞, by a similar argument to the one just used.

It follows that the total multiplicity of all the poles is 2, which must therefore be the Brouwer degree.

The only solution of $f(z) = \frac{1}{z} + z = 2$ is $z = 1$, and hence, by the Brouwer degree theorem, this point must have multiplicity 2. It *is* possible to calculate this directly by looking at the image of a small loop round $z = 1$ — for instance, $l(t) = 1 + ke^{it}$ $(0 \leq t \leq 2\pi)$ — but if you try this you may get bogged down in the algebra. Try it as a "challenge problem" if you like — it requires only techniques you have used already.

SAQ 8 Show that all zeros of $f(z) = z^4 + 2z + 4$ satisfy $1 < |z| < 2$.

SAQ 9 The function f from the Riemann sphere to itself maps the point with polar coordinates (r, θ) to the point with polar coordinates $(r, 3\theta)$ and $f(\infty) = \infty$; that is,

$$f\big(r(\cos\theta + i \sin\theta)\big) = r(\cos 3\theta + i \sin 3\theta).$$

What is the Brouwer degree of f?

SAQ 10 The function f from the Riemann sphere to itself is given by

$$f(z) = \begin{cases} z + \dfrac{z^3}{|z|}, & \text{if } z \neq 0, \infty, \\ 0, & \text{if } z = 0, \\ \infty, & \text{if } z = \infty. \end{cases}$$

By finding the multiplicity of the poles, show that the Brouwer degree is 3.

SOLUTION TO SAQ 8 We know that $|z^4| > |2z + 4|$ on $|z| = 2$ (see SAQ 7). Therefore $V_{l_2}(z^4) = V_{l_2}(z^4 + 2z + 4)$, where l_2 is the circle $|z| = 2$ taken once positively. But $z^4 = 0$ has a zero of multiplicity 4 at $z = 0$; so $V_{l_2}(z^4) = 8\pi$, and hence $z^4 + 2z + 4$ has four roots (which may all be different) inside l_2.

Again from SAQ 7, we know that $|z^4| < |2z + 4|$ on l_1 (the circle $|z| = 1$ taken once positively), and so

$$V_{l_1}(z^4 + 2z + 4) = V_{l_1}(2z + 4).$$

But $2z + 4 = 0$ has no roots inside l_1, and hence $V_{l_1}(2z + 4) = 0$. Hence $z^4 + 2z + 4$ has no roots inside l_1 either. Since all the roots lie inside l_2, they must lie in the annulus $1 \leq |z| < 2$; it is easy to see that there are no roots with $|z| = 1$, because if $|z| = 1$ then $|z^4 + 2z| \leq 3$, so $z^4 + 2z + 4 \neq 0$. ■

SOLUTION TO SAQ 9 The only point with $f(z) = 0$ is $z = 0$ and, taking a small loop $l(t) = ke^{it}$ $(0 \leq t \leq 2\pi)$ about 0, we see that the angular coordinate of $f(l(t))$ is just $3t$; so the winding number, and hence the Brouwer degree, is 3. ■

SOLUTION TO SAQ 10 The only pole of f is ∞ itself, and if we take a loop $l(t) = ke^{-it}$ $(0 \leq t \leq 2\pi$, k large$)$ which has winding number 1 about ∞, we see that $f(l(t)) = ke^{-it} + k^2(e^{-3it})$. Since $|k^2e^{-3it}| > |ke^{-it}|$ on l for large $k > 1$, we can apply Rouché's theorem: $f(l(t))$ has the same winding number about 0 as has $k^2(e^{-3it})$, i.e. -3; so the multiplicity of ∞ is $+3$, and so is the Brouwer degree. (The multiplicity of the zero at $z = 0$ is only 1, as you can check. This means that $f(z) = 0$ has at least one other root. In fact, it has two other roots, at $\pm i$.) ■

7.4 Applications of Winding Number

We have seen how to define multiplicity of a zero in general. This agrees with the simpler definition for polynomials to which you are used. If $f(z)$ is a (real or complex) polynomial with a zero at a, then $(z - a)$ is a factor of f. It may happen that $(z - a)^2$ or a higher power of $(z - a)$ divides $f(z)$. We can write

$$f(z) = (z - a)^r g(z),$$

where $r \geq 1$ and $(z - a)$ does *not* divide $g(z)$, so that $g(a) \neq 0$. The number r is called the *algebraic multiplicity* of the zero a. In this section, **B** shows that this number is the same as the multiplicity defined earlier. As usual, he works with the extended complex numbers $\mathbb{C} \cup \{\infty\}$, homeomorphic to the sphere S^2.

More generally, he considers rational functions $f: S^2 \longrightarrow S^2$ which are the quotient of two polynomials $f(z) = p(z)/q(z)$. To make sense of this quotient when $q(z) = 0$, we say that if a is a root of $q(z)$ (and not a root of $p(z)$), then a is a *pole* of $f(z)$, setting $f(z) = \infty$. **B** defines the (algebraic) multiplicity of these rational functions, which again agrees with the definition of the previous section.

With this interpretation of multiplicity, the Brouwer degree theorem generalizes the well-known theorem of algebra that a polynomial of degree k has exactly k roots, provided multiplicities are counted. **B** also shows how Rouché's theorem can be used to locate approximately where the roots of polynomial equations lie.

READ B Section 4.4 from page 124 to page 126, line 17, and from page 127, line 18 to the end of Section 4.4.

NOTE *Page 127, lines* -3, -4
This remark is referred to later as "the extension of Rouché's theorem". ▲

EXAMPLE 6 **B** page 132, Question 2.

Solution

Set $f(z) = z^3 + z^2 + z + 4$. Then $f(0) = 4 > 0$, and $f(-2) = -2 < 0$, so the equation $f(z) = 0$ has a real root lying between 0 and -2.

We shall use the notation of **B**'s Example 3 and Figure 4.6 (**B** page 128). Again we shall consider the circle $|z| = 3$, but this is a coincidence — other examples may require a larger circle: it has to be large enough to contain all the roots of $f(z) = 0$. On this circle, we have

$$\begin{aligned} |z^2 + z + 4| &\leq |z|^2 + |z| + |4| \\ &\leq 9 + 3 + 4 \\ &< |z^3|, \end{aligned}$$

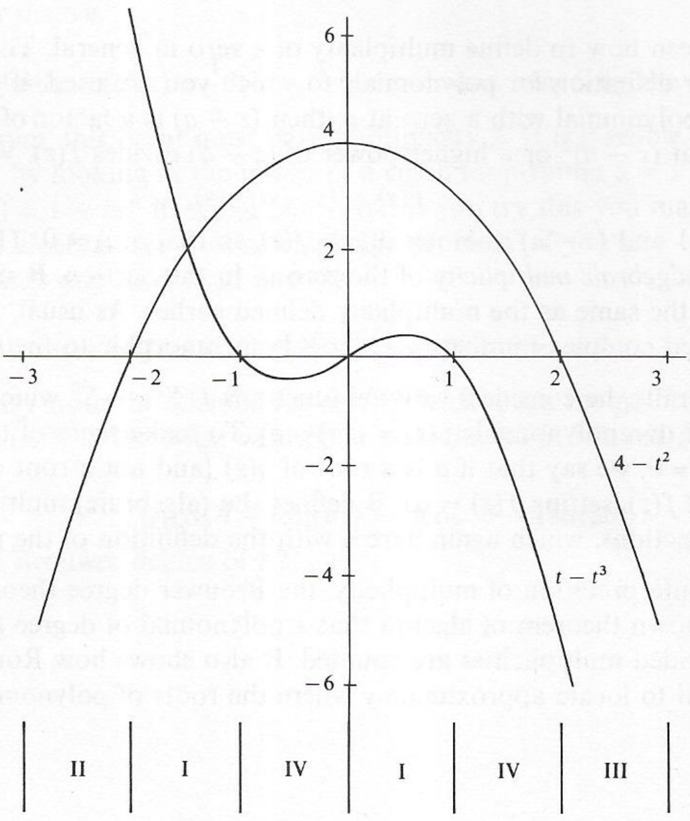
so the extension of Rouché's theorem applies to the quarter circle C_1, and we can conclude $|V_{C_1}(f(z)) - V_{C_1}(z^3)| < \pi$.

Now $V_{C_1}(z^3) = \dfrac{3\pi}{2}$, so $\dfrac{\pi}{2} < V_{C_1}(f(z)) < \dfrac{5\pi}{2}$. Exactly the same formula holds for C_2, C_3 and C_4.

On x_1, $f(z)$ is real, so $V_{x_1}(f(z)) = 0$ (and $V_{x_2}(f(z)) = 0$ similarly). On y_1, $z = it$, where t is real and $0 \leq t \leq 3$. So $f(z) = -it^3 - t^2 + it + 4$.

We can examine the variation of $\arg f(z)$ on y_1 by the technique used in SAQ 5 (page 18), so we sketch the real and imaginary parts of this expression. To deal with y_2 at the same time, we sketch values of

$$\left.\begin{aligned} 4 - t^2 &= \text{real part of } f(z) \\ t - t^3 &= \text{imaginary part of } f(z) \end{aligned}\right\} \text{ for } -3 \leq t \leq 3.$$



The edge y_1 starts at 0, so we can start with $\arg f(y_1(0)) = \arg f(4) = 0$. As t increases from 0 to 3, $y_1(t) = it$ and $f(y_1(t))$ moves from the first quadrant to quadrant IV and quadrant III. Hence $\arg f(y_1(3))$ lies between $-\frac{\pi}{2}$ and $-\pi$.

Similarly, the total rotation of the vector from 0 to $f(y_2(3))$ is from 0 through into quadrant IV, back to quadrant I, and into quadrant II, so is more than $\frac{\pi}{2}$ but less than π.

We can summarize our results so far as

$$V_{x_1} = 0, \quad \frac{\pi}{2} < V_{C_1} < \frac{5\pi}{2}, \quad -\pi < V_{y_1} < -\frac{\pi}{2},$$

$$V_{x_2} = 0, \quad \frac{\pi}{2} < V_{C_2} < \frac{5\pi}{2}, \quad \frac{\pi}{2} < V_{y_2} < \pi,$$

where the variation is understood to be of $f(z)$.

Now, the boundary B_1 of the area D_1 is $x_1 + C_1 - y_1$, so $V_{B_1} = V_{x_1} + V_{C_1} - V_{y_1}$. The inequalities give us

$$\frac{\pi}{2} < V_{C_1} < \frac{5\pi}{2}, \quad \frac{\pi}{2} < -V_{y_1} < \pi,$$

so, adding, $\pi < V_{B_1} < \frac{7\pi}{2}$.

Since V_{B_1} is an integral multiple of 2π, $V_{B_1} = 2\pi$, and so f has exactly one root in D_1.

You can check similarly that $\pi < V_{B_4} < \frac{7\pi}{2}$, so that f has one root in D_4. □

There are more examples in the Exercise Booklet of the use of Rouché's theorem in locating zeros of complex polynomials.

Appendix I (Non-assessed and optional)

In the tape section, we promised to supply the proof that we can cut $[a,b]$ into a finite number of pieces so that each piece is mapped by q into less than half of C. We now give the proof, which is not assessable. The proof will use the terminology of *M203* analysis, though you should be able to understand it without being familiar with *M203*.

The proof really says "suppose we cannot cut $[a,b]$ into a finite number of pieces with the desired property. Perhaps we can find a smaller interval $[a,c]$ which is OK?" Certainly we can get started, for if c is near enough to a then all the points $q(t)$ for $a \leq t \leq c$ will be near $q(a)$ — say less than halfway round C — since q is continuous. So $[a,c]$ is OK; it has the required property. But we can always extend an interval $[a,c]$ by adding on one more piece from c to c'. Again by continuity at c, we can make sure $q(c')$ and all the intermediate $q(t)$s are near to $q(c)$. But this extends our interval for a to c'. So perhaps we can get the whole way to b by repeating this process? In fact, we can.

The details of the proof are almost all in the preceding paragraph, but one needs a little care to give a complete proof.

Step 1 There is a number c, $a < c < b$ such that $[a,c]$ can be cut into a finite number of pieces, each of which q maps into less than half the circle.

Step 2 Let k be the least upper bound (see *M203* Section 8.3) of the cs satisfying the conditions of Step 1. If $k = b$, proceed to Step 4. If not, proceed to Step 3.

Step 3 If $k < b$, use the continuity of q at k to find a small interval $[k - d, k + d]$, containing k, so that this interval is mapped by q into less than half the circle.

Step 4 By assumption, $[a, k - \frac{1}{2}d]$ can be cut into a finite number of pieces, each of which maps into less than half the circle under q. Therefore if we add the piece $[k - \frac{1}{2}d, k + d]$, we can cut $[a, k + d]$ into a finite number of pieces with the desired property. But this contradicts the definition of k as a least upper bound, so in fact $k = b$ in Step 2.

Step 5 Now repeat Steps 3 and 4 with $k = b$: we deduce that $[a, b - \frac{1}{2}d]$ can be cut into a finite number of pieces with the right property, and now we need only add the piece $[b - \frac{1}{2}d, b]$ to these to complete the proof.

Perhaps we should have given a little more proof of Steps 1 and 3, but if you have understood the proof so far, you should be able to complete the details of these steps.

Appendix II (Non-assessed and optional)

Given a continuous loop l: $[a, b] \longrightarrow \mathbb{R}^2$, we can always find *continuous* functions $\theta(t)$, $r(t)$ so that the polar coordinates of $l(t)$ are $(r(t), \theta(t))$. We can deduce this easily from the work in the television programme.

Recall from the TV programme that we constructed a continuous map q from $[a, b]$ to the circle of directions. The polar coordinates of $l(t)$, (r, θ), are obtained quite easily from $q(t)$. The angle θ is just the angle the radius from 0 to $q(t)$ makes with the x-direction (and $r(t)$ is the distance from 0 to $l(t)$ along this radius). To ensure that we choose the right value for θ for each value of t to make $\theta(t)$ continuous, we have to use the function $\hat{q}$ from $[a, b]$ to the spiral. The height of $\hat{q}(t)$ above $q(a)$ tells us how far the radius from 0 to $q(t)$ has turned from its original position: as we said before, a turn of α corresponds to a change in height of $\frac{\alpha}{2\pi}$. So we take

$$\theta(t) = 2\pi\big(\text{height } \hat{q}(t) - \text{height } q(a)\big) + \theta(a),$$

where $\theta(a)$ is *any* value for the angular coordinate of $l(a)$. Then this is certainly a possible value for the angular coordinate of $l(t)$, and it is continuous since $q(t)$ is, as we know from the TV programme.

(**B** gives a different proof of this result in Section 4.3, which we do not want you to read.)

UNIT 8

POINCARE INDEX THEOREM REVISION MODULE

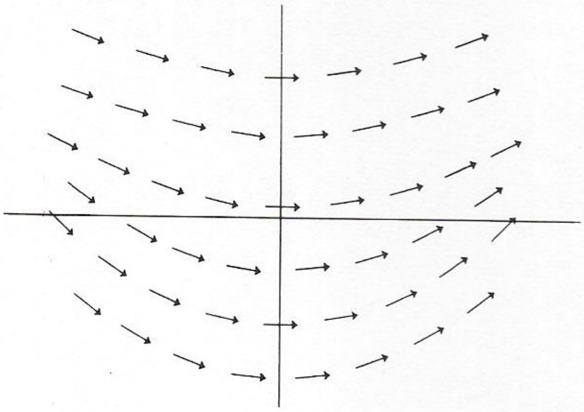
8.0 Introduction

This unit is rather different from the other units in the course and is divided into two parts. The second part is devoted entirely to a summary of the main ideas of the course, and is designed to help you with your revision. The first part explores the relation between the *topological* ideas we have been studying, and the theory of differential equations. It was after Henri Poincaré's work on differential equations that topology became a major subject in its own right. Today, the connection between differential equations and *topological manifolds* (the n-dimensional generalization of two-dimensional surfaces) is a very active area of research.

B deals briefly with some of the ideas of flows and differential equations in Chapter 5. Many of these ideas will be familiar to you from *M203*, and so we deal with them in rather more detail than **B**. However, they will be assessed only in the examination, and you will not be asked to reproduce any of the proofs. The examination questions will be broadly similar to SAQs 4, 6, 7 or 8.

8.1 Singularities of Flow Fields

If you have met the idea of a flow on the plane in *M203*, you may find it useful to refer to Section 2.2 of *Unit 2* and Section 19.3 of *Unit 19*. A *flow field*, also known as a *vector field*, on the plane is a function V which assigns to each point (x, y) of $\mathbb{R}^2$ a *vector* $V(x, y)$ based at that point. You should think of this vector as specifying a speed and a direction at the point.



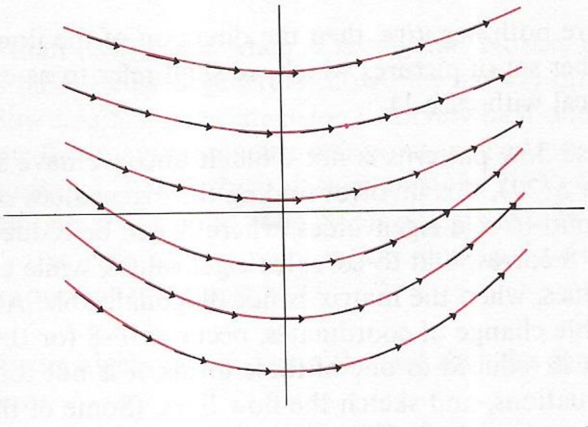
For example, if the plane represents a flat lake in a state of motion, the vector $V(x, y)$ at each point of the lake could be taken to represent the speed and direction with which the (idealized) particle of water is moving at that point.

Many other problems of applied physics, not just fluid motion, can be modelled using flow fields. In these problems, the function V usually satisfies some extra conditions. For instance, in the above example of a flat lake, we could expect that if the flow was smooth, with no turbulence, then the function V would be a smooth, or *differentiable* function. This means that if we look at the coordinates of $V(x, y)$,

$$V(x, y) = (f(x, y), g(x, y)),$$

then the functions f, g, which are functions from $\mathbb{R}^2$ to $\mathbb{R}$, will be differentiable; and, in fact, we shall assume from now on that the functions f and g have continuous partial derivatives with respect to x and y. This condition is satisfied in many applications to real problems.

The importance of this assumption is that it implies that we can find a *flow line* of the flow field through any point. This is a curve $\alpha:\mathbb{R}\longrightarrow\mathbb{R}^2$, whose tangent at any point gives the vector of the flow field at that point.



In other words, at each point $\alpha(t) = (\alpha_1(t), \alpha_2(t))$, the tangent $\alpha'(t) = (\alpha_1'(t), \alpha_2'(t))$ to the curve is just the vector $V(\alpha_1(t), \alpha_2(t))$ of the flow field.

The proof of this is really just a proof that the "differential equation" (or rather *pair* of equations) $(\alpha_1'(t), \alpha_2'(t)) = V(\alpha_1(t), \alpha_2(t))$ has a solution. A special case is the pair of equations

$$\frac{dx}{dt} = f(x, y), \quad \frac{dy}{dt} = g(x, y), \tag{1}$$

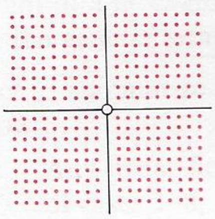
where we write $x(t)$ for the first coordinate, $\alpha_1(t)$, of α, $y(t)$ for the second, and f, g for the coordinates of the vector $V(x, y)$. Here, we have assumed that the right-hand sides are independent of t; that is, the flow field itself is fixed and does not change as t changes. The parameter t often represents time, and this restriction to flow fields which do not change with time leaves many cases of interest.

The simplest cases to look at are *linear flows* where the functions f, g in (1) are *linear* functions: $f(x, y) = ax + by$, $g(x, y) = cx + dy$. So, in a given coordinate system, a linear flow is described by an equation

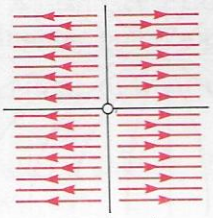
$$\begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{A} \begin{pmatrix} x \\ y \end{pmatrix},$$

where $\mathbf{A}$ stands for the 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

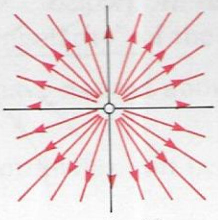
These equations can be solved explicitly. This was done for many cases in *M203*; the result is that the pattern of flow lines must resemble one of the following pictures. In this list, a and b are *positive* real numbers.



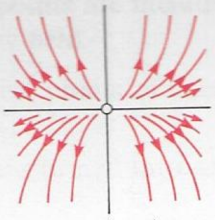
1 $\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ (zero flow)



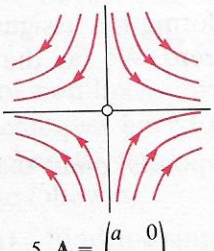
2 $\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$



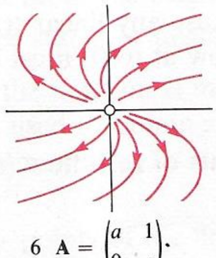
3 $\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$



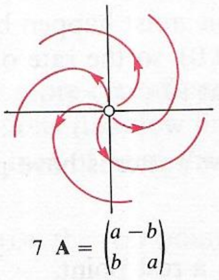
4 $\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$



5 $\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & -b \end{pmatrix}$

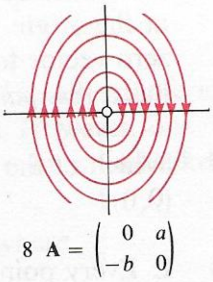


6 $\mathbf{A} = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$.



7 $\mathbf{A} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$

$(a^2 + b^2 = 1)$



8 $\mathbf{A} = \begin{pmatrix} 0 & a \\ -b & 0 \end{pmatrix}$

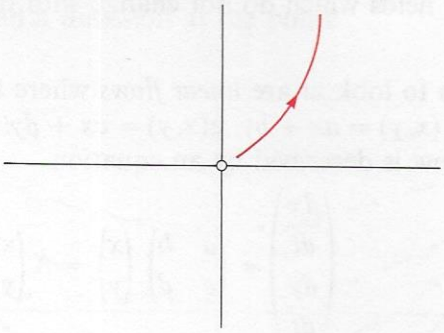
If a and b are both *negative*, then the direction of the flow is *reversed*. This gives rise to another set of pictures which we shall refer to as cases $2', \ldots, 8'$ (case $1'$ being identical with case 1).

Plotting these flow patterns is not difficult once we have solved the equations. As shown in *M203*, this involves finding the eigenvalues of the matrix **A**. Cases 1–5 correspond to real eigenvalues, where **A** can be reduced to diagonal form. Cases 7 and 8 correspond to complex eigenvalues, while case 6 corresponds to real eigenvalues, when the matrix is not diagonalizable. All linear flow patterns, after a suitable change of coordinates, occur as 1–8 (or their "reverses"). Once the matrix **A** is reduced to one of these forms, it is not too difficult to solve the system of equations, and sketch the flow lines. (Some of the details may be found in the tape section of *M203 Unit 19*.)

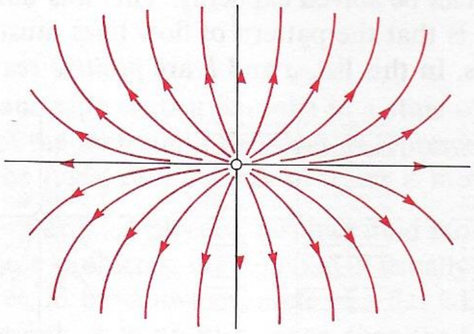
EXAMPLE 1 Consider case 4. The solutions of

$$\begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{are} \quad \begin{cases} x = c_1 e^{at}, \\ y = c_2 e^{bt}, \end{cases}$$

and we can plot the flow line through any point by plotting those curves for suitable values of c_1, c_2. For instance, to find the flow line through $(1, 1)$ take $c_1 = c_2 = 1$. This passes through $(1, 1)$ at $t = 0$. Now calculating values of e^{at}, e^{bt} for other values of t, we obtain the curve shown below.



Plotting lines for other values of c_1 and c_2 (positive and negative) gives the flow pattern of type 4.



□

A number of features are worth pointing out. In each example, the flow is zero at the origin. This must happen because any linear transformation assigns the zero vector to $(0, 0)$, so the rate of flow at $(0, 0)$ must be zero. We say the flow has a *rest point* at $(0, 0)$.

SAQ 1 Which of the flow patterns have points of zero flow (rest points) *other* than $(0, 0)$?

SOLUTION TO SAQ 1

1. Every point is a rest point.
2. Every point on the y-axis is a rest point. ■

Rest points other than $(0, 0)$ occur when $\mathbf{A}$ is *singular*. Notice, too, that a rest point P is also an (admittedly degenerate) flow line, namely $\alpha(t) = P$ for all $t \in \mathbb{R}$. Any other flow line is a one-dimensional curve which does *not* contain a rest point. The flow line may *approach* a rest point, or be traced back arbitrarily close to one as in cases 2, 3, 4, 6, 7 (or dashed versions 2′, 3′, 4′, 6′, 7′); but it cannot actually contain a rest point. The closer a particle is to a rest point, the slower it moves! So, in case 3′, as a point moves along a flow line towards O, its speed decreases to prevent it arriving at O.

SAQ 2 Check that this is true algebraically, using the fact that the general solution of the pair of equations

$$\begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \end{pmatrix} = \begin{pmatrix} -a & 0 \\ 0 & -a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

is $x = c_1 e^{-at}$, $y = c_2 e^{-at}$.

SOLUTION TO SAQ 2 Since the solutions are of the form ke^{-at}, as t grows the magnitudes of both x and y decrease—i.e. any point of a solution curve moves towards O. However, since $e^x \neq 0$ (any x), we never have $(c_1 e^{-at}, c_2 e^{-at}) = (0, 0)$ unless $c_1 = c_2 = 0$. ■

If we interpret 3, or 3′, as models of a fluid flow, for example a flat lake, then, at a rest point, the fluid is stationary, and other water particles flow towards or away from the rest point. This would imply a compression of the fluid or, if the fluid is not very compressible, like water, a situation where fluid was actually being removed from or added to the system at the rest point. Hence these rest points are often called "sinks" and "sources", respectively.

We can split the pictures into two classes in the following way. If a point very close to the rest point flows away "to infinity", then we say that the rest point is *unstable*. A small perturbation leads, ultimately, to vastly different behaviour. If, on the other hand, any point close to the rest point remains close to it, then we say that the rest point is *stable*.

Thus group 1 (stable): 1, 2′, 3′, 4′, 6′, 7′, 8, 8′;
group 2 (unstable): 2, 3, 4, 5 (=5′), 6, 7.

Another term frequently used is stable or unstable *equilibrium points*; this is because in some applications, a rest point of a flow represents an equilibrium point of a mechanical system (see Example 2 below). *All* rest points are included in the general term "singular point" or "singularity" of the flow field.

Let us summarize something of what we have noticed about rest points of linear flows.

1. Any linear flow, after a change of coordinates, can be transformed into one of cases 1–8, 2′–8′.
2. Every linear flow apart from the zero flow has either a line of rest points or just one rest point at the origin.
3. Rest points can be stable or unstable.

For non-linear flows, the position is much more complicated. Usually, numerical techniques have to be used to obtain solutions (the flow lines). However, one technique which sometimes gives us some idea of what the flow lines look like is the following.

(a) Find the singular points of the flow (i.e. the rest points, where $\dfrac{dx}{dt} = \dfrac{dy}{dt} = 0\Big)$.

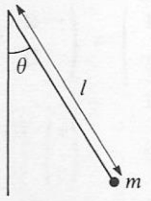
(b) Near each singular point, replace the actual flow by the linear approximation to it. That is, if the flow is given by (1), we look at the linear flow where f and g are replaced by their linear (first-order) Taylor approximations at the singular point.

(c) Determine the nature of the linear flows obtained.

This process sometimes gives us a good idea of what the non-linear flow looks like, but it must be stressed that in some cases step (b) can drastically alter the position.

EXAMPLE 2 The equation of motion of a pendulum consisting of a mass at the end of a light inextensible string (or rod) freely fixed at one end, and ignoring air resistance, is

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\sin\theta,$$



where l is the length of the string. This can be written in our form by putting $\omega = \frac{d\theta}{dt}$; then we obtain a system of equations

$$\frac{d\theta}{dt} = \omega, \quad \frac{d\omega}{dt} = -\frac{g}{l}\sin\theta.$$

This is not linear, but the linear approximation is

$$\frac{d\theta}{dt} = \omega, \quad \frac{d\omega}{dt} = -\frac{g}{l}\theta.$$

The point $\theta = \omega = 0$ is a stable rest point of both systems, and corresponds to the "stable equilibrium" of the weight when it is hanging vertically down.

The second system is widely used for modelling the motion of a pendulum, and in fact gives a good approximation for small angles θ. The matrix **A** is non-singular and corresponds to one of our standard examples (case 8). The flow lines are ellipses centred at the singular point.

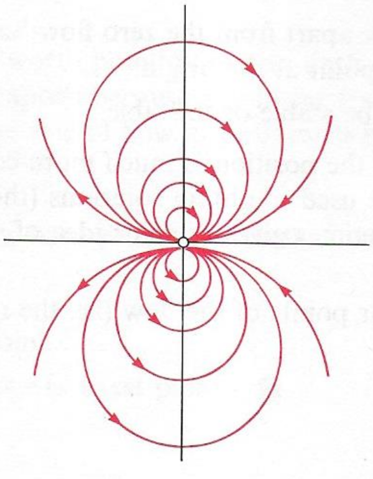
EXAMPLE 3 Sometimes we can solve the equations exactly for a non-linear flow. For example, look at the system

$$\frac{dx}{dt} = x^2 - y^2, \quad \frac{dy}{dt} = 2xy.$$

If we write $z = x + iy$, the above system can be summarized as

$$\frac{dz}{dt} = z^2$$

which can certainly be solved explicitly. The flow lines turn out to resemble those of a magnetic "dipole".



This is an example where our proposed technique would not work: the linear approximations to $x^2 - y^2$ and $2xy$ at $(0,0)$ are, in fact, both the zero function, so the "best" linear flow would be the plane of rest points (case 1). This is not a very good approximation!

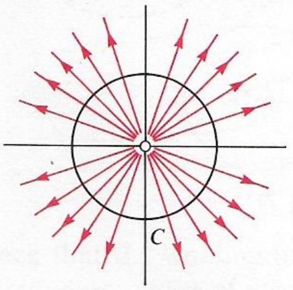
And if you look carefully at the diagram above and compare it with our case list, you will see it is not like any of them near $(0,0)$: the trajectories approach $(0,0)$ in the halfplane $x > 0$ and move away from $(0,0)$ in the halfplane $x < 0$. □

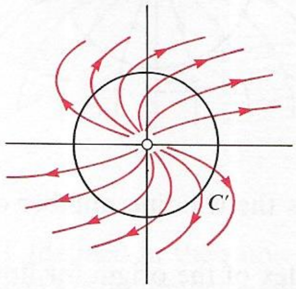
By now it should seem plausible that topologists have an interest in flows. Even when an exact analytical solution cannot be found, the "shape" of the flow near a singular point can provide valuable information of the nature of the equilibrium there. We have already used the fact that a general linear flow may be transformed into one of our diagrams by a change of coordinates; we now allow ourselves to use *any homeomorphism*, not just a linear change of coordinates. If two flows can be transformed into one another by homeomorphisms, we are justified in saying they have the "same shape", and call the flows "equivalent".

Q 3 Which of cases 4–8 is equivalent to case 3?

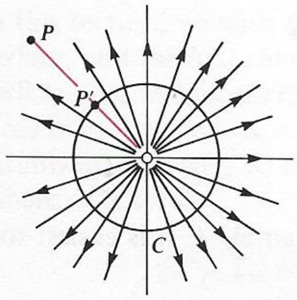
ON
Q 3 It is fairly clear that case 4 is equivalent to case 3: we can imagine simply "straightening out" the flow lines. Not quite so obvious is the fact that 6 and 7 are also equivalent to 3. ■

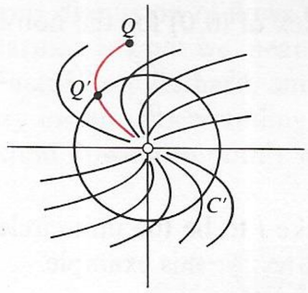
To indicate why this rather surprising fact is true, we shall show how to construct a smooth homeomorphism of flow 3 with flow 6.





Start with a small closed curve C around the origin which meets each flow line of flow 3 exactly once. Do the same thing with a curve C' round the origin for flow 6. In these examples, a circle centre $(0,0)$ will do. Now choose any smooth homeomorphism between C and C'—we could take C and C' to be the same circle and use the identity homeomorphism. Extend this to points outside C as follows. Each point P lies on a *unique* flow line, and at time t_P earlier P lay on C (at a point P').





Now take the corresponding point Q' on C', and let it flow for time t_P. The point Q thus obtained is the image of P. Since any point inside C (except $(0,0)$ itself) flows to some point of C, we can use the same trick (backwards) inside C everywhere except $(0,0)$. Finally, we define the image of $(0,0)$ to be $(0,0)$.

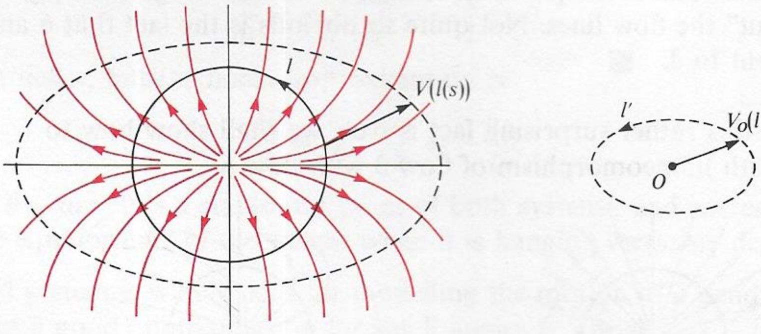
It is very plausible that this defines a homeomorphism of $\mathbb{R}^2$ to itself, but a formal proof would require some theorems concerning solutions of differential equations. A similar proof would show that case 7 is also equivalent to these two, but a little more care is required to find a suitable curve C'. Of the remaining cases, 5 and 5′ and 8 and 8′ are equivalent, but 1, 2 and 2′ are equivalent only to themselves.

Taking account of the reverses, we end up with the following list of equivalent cases:

$$\{1\}, \{2\}, \{2'\}, \{3, 4, 6, 7\}, \{3', 4', 6', 7'\}, \{5, 5'\}, \{8, 8'\}.$$

In fact, there is a very simple way of extracting some of the topological information from a flow picture. Given any *isolated* singular point P of a flow (not necessarily linear), we can define an integer called the *index* of the singular point, as follows. Draw a small simple loop l round P which winds once positively round P. (For the time being, "small' just means small enough to exclude any other singular points from its interior.) Now, at each point $l(s)$ of the loop, there is a *non-zero* flow vector $V(l(s))$.

If we consider a corresponding vector $V_O(l(s))$, equal to $V(l(s))$ but based at a point O, the "sharp end" of $V_O(l(s))$ will trace out a loop l' as $l(s)$ traces out l. Since $V(l(s))$ is never zero on l, neither is $V_O(l(s))$; so l' does not pass through O.



DEFINITION The *index* of P is the winding number of l' about O.

EXAMPLE 4 Compute the index of the origin for flow 4.

Solution

Take l to be the circle $[0, 2\pi] \longrightarrow \mathbb{R}^2$ given by $s \longmapsto (\cos s, \sin s)$. The matrix $\mathbf{A}$ of the flow is $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, so the vector $V_O(l(s))$ is $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} \cos s \\ \sin s \end{pmatrix} = \begin{pmatrix} a\cos s \\ b\sin s \end{pmatrix}$. Hence $l'(s) = (a\cos s, b\sin s)$, and the winding number about $(0, 0)$ is $+1$, which is the index we seek.

EXAMPLE 5 Compute the index of $(0, 0)$ for the non-linear flow

$$\frac{dx}{dt} = x^2 - y^2, \quad \frac{dy}{dt} = 2xy.$$

Solution

Again, we can take l to be the unit circle as in Example 4. It is easiest to work in polar coordinates for this example.

The vector at $l(s) = (\cos s, \sin s)$ is $(\cos^2 s - \sin^2 s, 2\cos s \sin s)$, or $(\cos 2s, \sin 2s)$. Hence the polar coordinates of $V_O(l(s))$ are $[1, 2s]$, and the winding number of the loop l' is therefore 2, which is the index of $(0, 0)$.

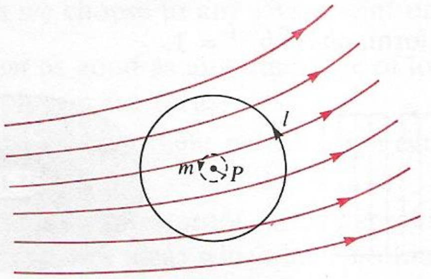
SAQ 4 Compute the index of flows 3′ and 5.

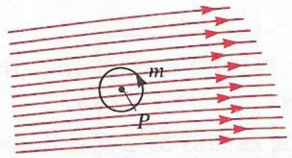
OLUTION TO SAQ 4 (i) 3′. Working as before, if $l(s)$ is the point $(\cos s, \sin s)$, then $V_O(l(s))$ is $\begin{pmatrix}-a & 0\\ 0 & -a\end{pmatrix}\begin{pmatrix}\cos s\\ \sin s\end{pmatrix}$, so the polar coordinates of $l'(s) = V_O(l(s))$ are $(a, s+\pi)$ and the winding number, or index, is 1. This may come as a shock, since instinct suggests that reversing the flow might change the sign of the index. However, this is not the case. The loops l' for cases 3 and 3′ are both traced once positively; in case 3′, the starting point is $(-a, 0)$, whereas in case 3 it is $(a, 0)$.

(ii) 5. We find
$$l'(s) = V_O(l(s)) = (a\cos s, -b\sin s)$$
$$= (a\cos(-s), b\sin(-s)),$$
and this time the index is -1. ■

In the next section, we shall put the index to good use. First, we must clear up two points. It may have occurred to you that the definition of index depends on the choice of simple loop l. The proof that it does not uses the fact that winding numbers are not changed by a homotopy: any two simple loops l_1, l_2 are homotopic, and can be deformed into one another, and the homotopy gives us a homotopy of l'_1 with l'_2. We shall say no more about this proof.

Lastly, we make an observation which will be useful in the next section. Although we have defined the index only of a singular point of a flow, the definition extends easily to a non-singular point. The fact is that the "index" of a non-singular point must be zero. To see this, choose as before a simple loop surrounding the point. Now by a homotopy we can shrink the loop to a very small one: so small that the flow in and around it is almost constant.



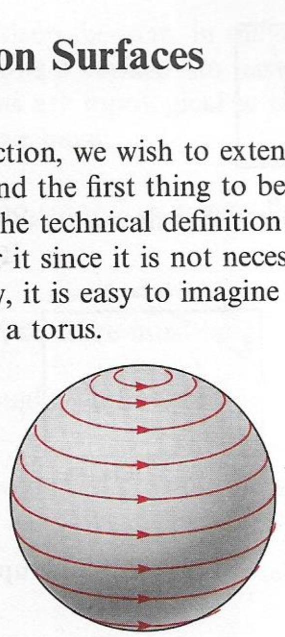


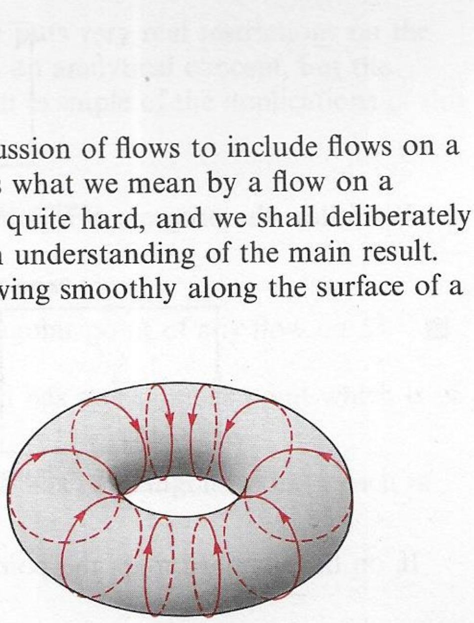
SAQ 5 Check that the winding number of l' (defined in the same manner as the l' used to define the index of a singular point) for an almost constant flow is zero.

SOLUTION TO SAQ 5 For an almost constant flow, $l'(s)$ varies only a very small amount from a fixed value, and so l' does not wind round the origin. ■

8.2 Flows on Surfaces

In this section, we wish to extend our discussion of flows to include flows on a surface, and the first thing to be clarified is what we mean by a flow on a surface. The technical definition is actually quite hard, and we shall deliberately gloss over it since it is not necessary for an understanding of the main result. Intuitively, it is easy to imagine a fluid flowing smoothly along the surface of a sphere or a torus.





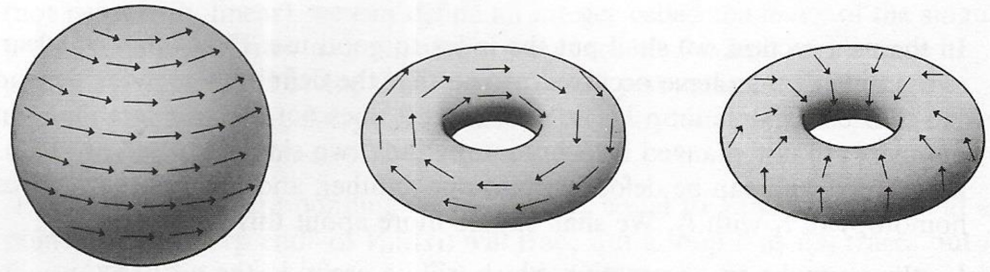
(Alternatively, imagine the surface growing hair! If the hair is combed to lie along the surface, then a small segment of hair near the root gives a "flow field" on the surface.)

On the other hand, it is not immediately obvious what a smooth flow on a tetrahedron looks like. What happens on the edges and at the vertices?

To dodge this problem, we shall henceforth assume that all our surfaces are "smooth". (This is no great restriction, since a surface assembled from polygons can always be deformed into a "smooth" surface.)

A *flow field* on a surface S is a function which assigns to each point P of S a tangent vector to S at P.

EXAMPLE 6

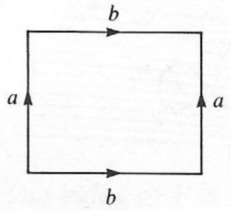


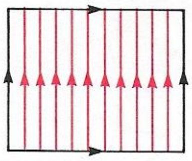
□

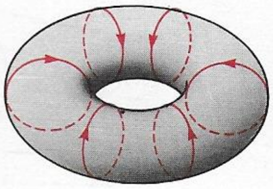
The problem of analysing a general flow on a general surface is very hard, because it amounts to solving a non-linear system of differential equations. Here is an example of how delicate things can be.

EXAMPLE 7 (Flows on a torus)

Consider the torus in its canonical form, $aba^{-1}b^{-1} = 1$.

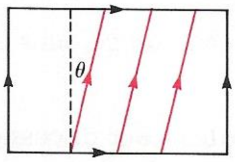




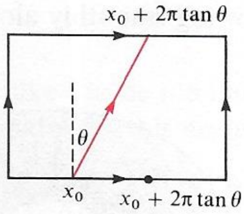


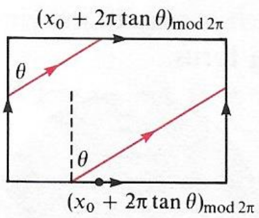
We can define a flow on the torus simply by drawing vertical arrows on the above picture; as shown, when the identifications are performed, this leads to "meridian" flow. Every flow line is a closed curve.

Now alter the flow slightly, so that on the standard representation the flow lines are inclined at an angle θ to the vertical.



We follow the progress of a point on one particular flow line, say γ.




$(x_0 + 2\pi \tan \theta)_{\mathrm{mod}\,2\pi}$
θ
θ
$(x_0 + 2\pi \tan \theta)_{\mathrm{mod}\,2\pi}$

If, on the standard map, the point leaves $\left(x_0, -\frac{\pi}{2}\right)$, it will flow up to the point

$\left((x_0 + 2\pi \tan\theta)_{\text{mod } 2\pi}, \dfrac{3\pi}{2}\right)$ and re-appear on the bottom edge at the identified point $\left((x_0 + 2\pi \tan\theta)_{\text{mod } 2\pi}, -\dfrac{\pi}{2}\right)$. Thus the magnitude of the distance (measured along the torus) between successive crossings of the line $\phi = -\dfrac{\pi}{2}$ is given by

$$(2\pi \tan\theta)_{\text{mod } 2\pi}.$$

In exactly the same way, the distance from $\left(x_0, -\dfrac{\pi}{2}\right)$ to the nth crossing of $\phi = -\dfrac{\pi}{2}$ is given by

$$(2\pi n \tan\theta)_{\text{mod } 2\pi}.$$

There are now two distinct possibilities.

(i) $\tan\theta = \dfrac{p}{q}$ (a rational number). In this case, the qth crossing of $\phi = -\dfrac{\pi}{2}$ is at $\left(x_0, -\dfrac{\pi}{2}\right)$, and the solution curve is periodic thereafter, giving a closed flow line.

(ii) If $\tan\theta$ is not rational, then, by varying n, we can obtain a value of $(2\pi n \tan\theta)_{\text{mod } 2\pi}$ *as close as we choose* to any pre-assigned number in $[0, 2\pi)$.* This means that the flow line γ crosses the line $\phi = -\dfrac{\pi}{2}$ as close as we choose to any given point on it! Finally, since the line $\phi = -\dfrac{\pi}{2}$ is just as good as any other line of longitude, γ passes arbitrarily close to any point on the torus!

Thus a very slight perturbation can have catastrophic effects on the nature of the flow.

*This requires proof, strictly speaking, but we shall not give one as the proof uses ideas which have nothing to do with this course. □

It may appear that the analysis of flows on a surface is a hopeless task. However, there is one bold assertion we can make.

THEOREM 1 (Poincaré)
Let S be any closed orientable surface, and V a flow on S with only isolated singularities. Then the sum of the indices of the singular points of V is $\chi(S)$.

This is a deep theorem, in the sense that it puts very real restrictions on the type of flows a surface can have; flows are an analytical concept, but the restrictions are topological in nature. As an example of the applications of this theorem we have:

ROLLARY 1 There is no smooth non-zero flow on the sphere.

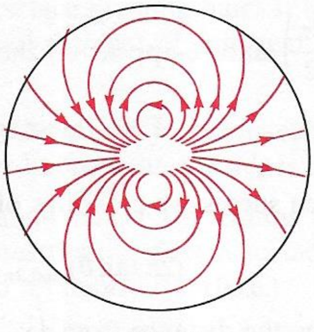
Proof

$\chi(S^2) = 2$, so there must be at least one singular point of any flow on S^2. ■

SAQ 6 Draw (roughly) a flow on the sphere which has one singular point which is of index 2.

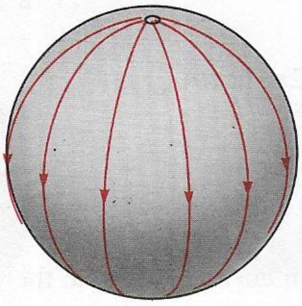
SAQ 7 Draw (roughly) a flow on the sphere which has two singular points, each of index 1.

SOLUTION TO SAQ 6



a "dipole" ■

SOLUTION TO SAQ 7



source at the north pole, sink at the south pole ■

COROLLARY 2 The torus is the *only* closed orientable surface which has a smooth flow without singular points.

Proof

For a closed orientable surface, $\chi = 2 - 2g$, whence $g = 1$ and the surface is a torus. ■

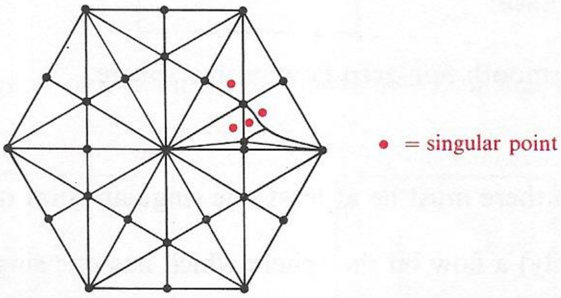
We end the formal part of the course with an outline of the proof of Poincaré's theorem. Although it is non-assessable, we hope you will read through it at least once, because the ideas are quite elegant.

Proof of Theorem 1

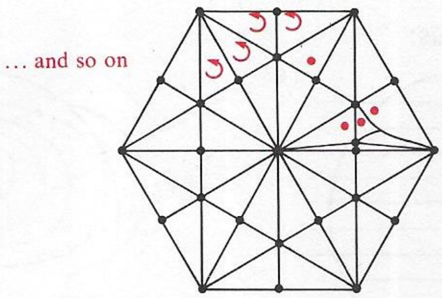
The proof is divided into two parts.

(i) We first prove that the sum of the indices of singular points of a flow V on S is independent of V—that is, *any* two flows V, W give rise to the same index sum.

To see this, consider S in its canonical form as a polygon with edge identifications. Now, the singular points of V are isolated, and so each can be surrounded by an open disc of non-singular points. Because S is compact, there are therefore only a finite number of singular points. Now dissect S into triangles in such a way that all the singular points of V and W lie strictly inside the triangles—not on an edge nor on a vertex—and no triangle contains more than one singular point of each.

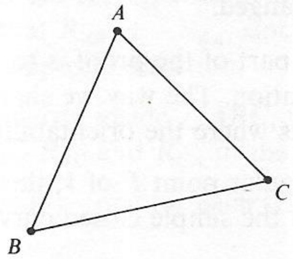


(This is always possible, because we can certainly "bend" a triangulation of S smoothly to avoid a finite number of points, and by subdividing we can make triangles small enough to enclose no more than one singular point.) Since S is orientable, we can choose a definite orientation of each triangle in such a way that the result is coherent.

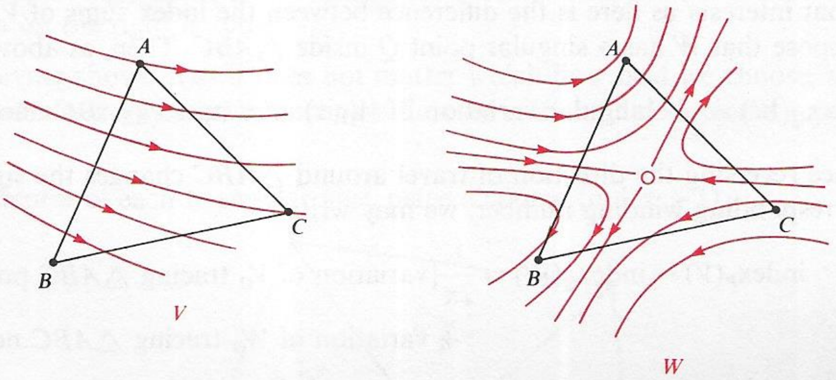


Thus each edge occurs in two adjacent triangles, from which it receives opposite orientations.

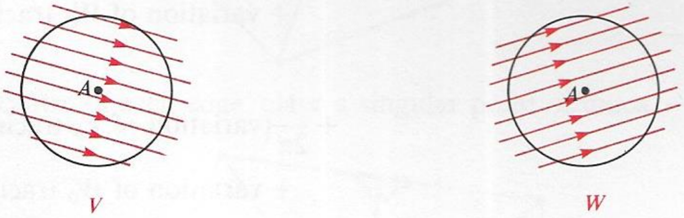
Now consider a typical triangle ABC.



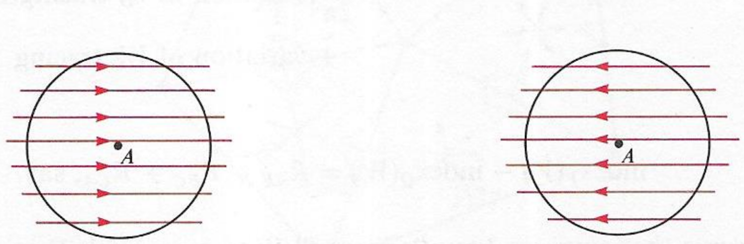
There are two flows, V, W, across $\triangle ABC$, and these are considered as quite separate entities.



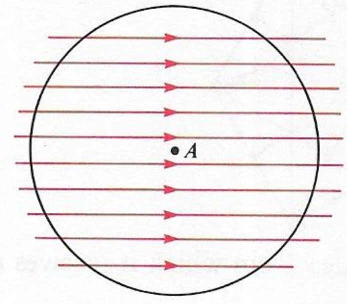
We shall deform W very slightly near the boundary of $\triangle ABC$ (and well away from any interior singular points), so that V and W agree at all three vertices. This is possible because in a *small* neighbourhood of a vertex, there are no singularities and the flow is "locally uniform".



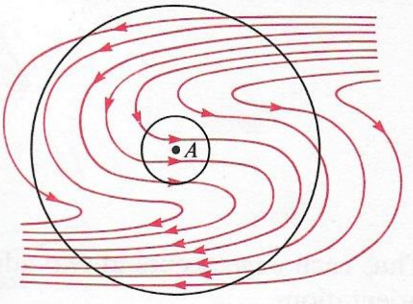
The worst extreme, when V and W have opposite directions at A, is shown below,



and is dealt with thus:



V



adjusted W

Making this alteration to W does not introduce any singular points, and so the index sum of W is unchanged.

The central idea of this part of the proof is to attach a number R_{AB} to each edge AB of the triangulation. The way we shall do this will make it clear that $R_{AB} = -R_{BA}$ (and this is where the orientability of the surface will be essential).

If $\triangle ABC$ encloses a singular point P of V, then one way to compute the index of P is to use $\triangle ABC$ as the simple closed curve l in the definition of index. Then

$$\text{index}_P(V) = \frac{1}{2\pi}\{\text{angular variation of } V_O(x) \text{ as } x \text{ traces } \triangle ABC \text{ once positively}\}.$$

What interests us here is the difference between the index sums of V and W, so suppose that W has a singular point Q inside $\triangle ABC$. Then, as above,

$$\text{index}_Q(W) = \frac{1}{2\pi}\{\text{angular variation of } W_O(x) \text{ as } x \text{ traces } \triangle ABC \text{ once positively}\}.$$

Since reversing the direction of travel around $\triangle ABC$ changes the sign of the corresponding winding number, we may write

$$\begin{aligned}\text{index}_P(V) - \text{index}_Q(W) = \frac{1}{2\pi}\{&\text{variation of } V_O \text{ tracing } \triangle ABC \text{ positively}\\ &+ \text{variation of } W_O \text{ tracing } \triangle ABC \text{ negatively}\}.\end{aligned}$$

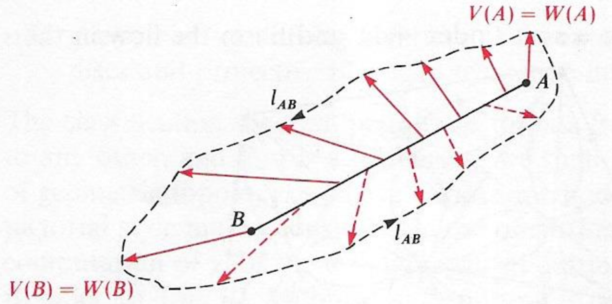
This expression may be split into three parts, one for each side of $\triangle ABC$:

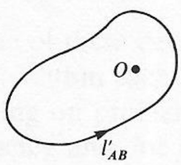
$$\begin{aligned}\text{index}_P(V) - \text{index}_Q(W) = &\frac{1}{2\pi}\{\text{variation of } V_O \text{ tracing } AB\\ &\qquad + \text{variation of } W_O \text{ tracing } BA\}\\ &+\frac{1}{2\pi}\{\text{variation of } V_O \text{ tracing } BC\\ &\qquad + \text{variation of } W_O \text{ tracing } CB\}\\ &+\frac{1}{2\pi}\{\text{variation of } V_O \text{ tracing } CA\\ &\qquad + \text{variation of } W_O \text{ tracing } AC\}\end{aligned}$$

i.e.

$$\text{index}_P(V) - \text{index}_Q(W) = R_{AB} + R_{BC} + R_{CA}, \text{ say.} \qquad (*)$$

Now observe that since we have "adjusted" W to agree with V on all vertices, R_{AB} is itself a winding number.





The curve l_{AB} is formed by tracing the sharp end of $V(x)$ as x moves along AB, and then the sharp end of $W(x)$ as x moves back to A. (The initial adjustment of W is precisely to make l_{AB} a loop rather than a pair of arcs.)

R_{AB} is the winding number of the corresponding loop l'_{AB} about our reference point. And this makes clear that $R_{AB} = -R_{BA}$, since the latter is obtained from the same curve traced in the opposite direction.

Finally, if V (W) has no singularity inside $\triangle ABC$, then choose P (Q) arbitrarily in the interior, and define R_{AB}, R_{BC} and R_{CA} in the same way.

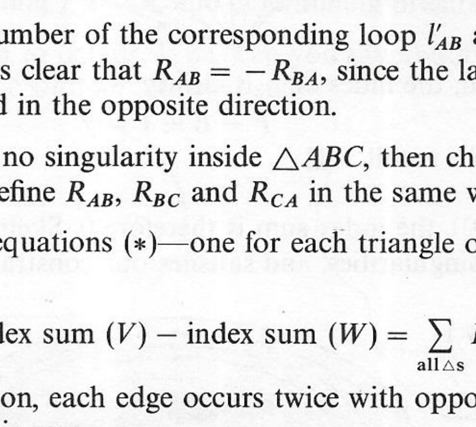
Now add up all the equations $(*)$—one for each triangle of the triangulation. We obtain

$$\text{index sum } (V) - \text{index sum } (W) = \sum_{\text{all} \triangle\text{s}} R_{AB}.$$

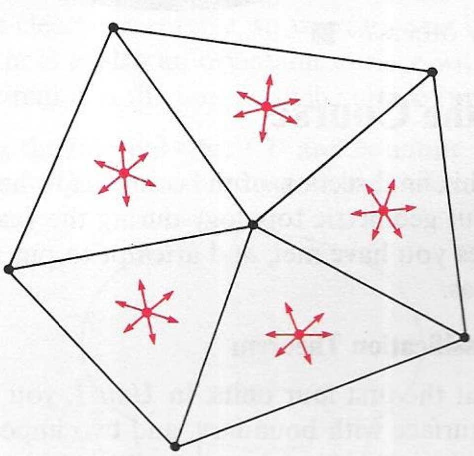
But in the triangulation, each edge occurs twice with opposite orientations, so the sum on the right is zero.

(ii) Having shown that it does not matter which flow field we choose, we now give a standard way of constructing a flow on a triangulated, closed orientable surface.

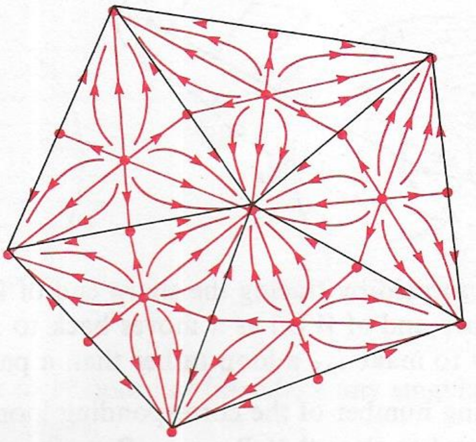
At the centre of each triangular face, place a source (index $+1$).



Now, at the centre of each edge, place a singular point of index -1, as follows.



Finally, at each vertex place a sink (index $+1$), and fill in the flow in the obvious way.



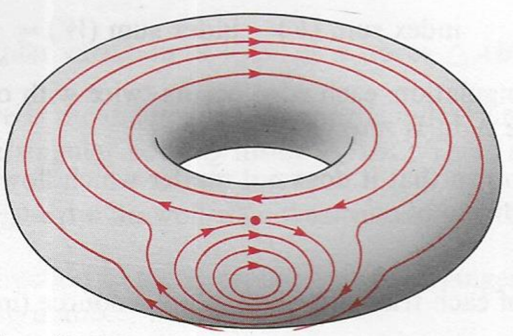
For this flow field, the index sum is clearly

$$F - E + V = \chi,$$

and this proves the result. ■

SAQ 8 On a torus ($\chi = 0$), the index sum is therefore 0. Sketch (roughly) a flow on the torus which has singularities, and satisfies our constraint.

SOLUTION TO SAQ 8 One possible flow is



Can you find any others? ■

8.3 Revision of the Course

The purpose of this final section of the course is to help you revise what you have learned about geometric topology during the year. We shall discuss various techniques you have met, and attempt to put into perspective the main mathematical ideas.

Part I: The Classification Theorem

Let us look first at the first four units. In *Unit 1*, you met the definition of a surface and of a surface with boundary, and two important topological invariants: Euler characteristic (χ) and number of boundary components (β). In *Unit 2*, we introduced the third (and final) invariant needed for the topological classification of surfaces, namely orientability. (*Unit 2* also dealt with regular subdivisions, but we shall not revise that here.)

In *Unit 3* we broached the principal technique: topological polygons with edge identifications and their accompanying edge equations. This led us to the idea of rules for manipulating edge equations, and then, in *Unit 4*, to the concept of the canonical equation for a surface. The classification theorem for surfaces was proved by showing that χ, β and orientability are not altered by the manipulation of edge equations, and can be read off from a canonical equation.

The classification theorem is *the* basic result in the course. In this revision section, we shall illustrate and employ it in three different ways:

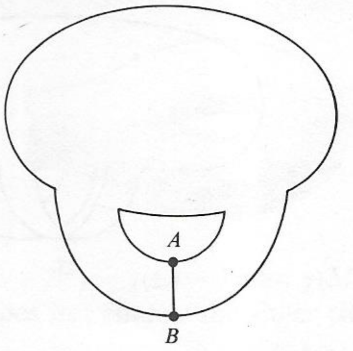
(i) *pictorial*—classifying a surface given a possibly unusual picture of it;

(ii) *algebraic*—classifying a surface given a possibly unusual edge equation for it;

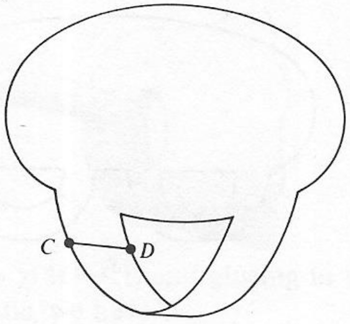
(iii) *geometric*—classifying a surface as a connected sum of a sphere with tori, discs and projective planes (a technique introduced in *Unit 4*).

The classification theorem permits us to pass from any one of these descriptions to any other, and that is a further use we shall have for it. Within each "style" of geometric topology, we have various methods for working on problems. The pictorial style makes considerable use of cutting and regluing and, for the computation of χ, of the technique called cutting handles (see the Exercise Booklet for *Unit 4*). Möbius band spotting is not so easy in this style, but is almost mechanical in the algebraic style—one looks for an edge equation of the form $AaBaC = 1$. The computation of χ in this style is the routine but tricky "insertion of vertices" technique. Lastly, the connected sum formalism has its own rules for calculating χ and β, and determining orientability.

It is time to get down to details. What can you say about each of these surfaces (each represents a disc with an attached strip)?



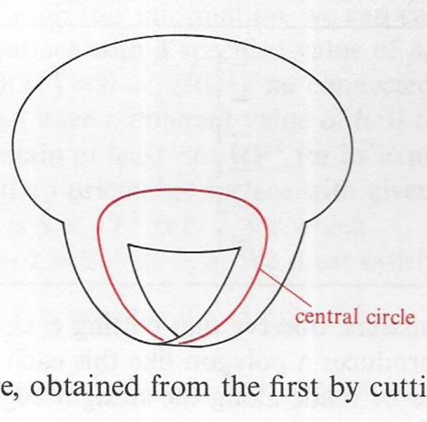
S_1



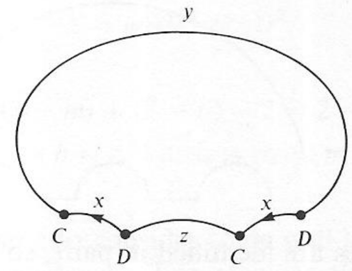
S_2

S_1 is clearly a disc with a hole, but if you were asked to prove this, one technique would be to introduce an internal edge, such as AB. The figure now has two vertices, so $V = 2$; 3 edges (a loop based at A, another at B, and the edge AB), so $E = 3$; and $F = 1$: $\chi = V - E + F = 0$. There are two boundary curves, so $\beta = 2$. It is clearly orientable, so we recognize it as a disc with a hole, since the disc with a hole is also an orientable surface with $\chi = 0$, $\beta = 2$, and, by the classification theorem it is the unique such surface (up to homeomorphism).

What is S_2? Marking the internal edge CD and counting, we find $V = 2$, $E = 3$, $F = 1$, so $\chi = 0$, but now $\beta = 1$, and moreover the surface is not orientable. Here is a Möbius band

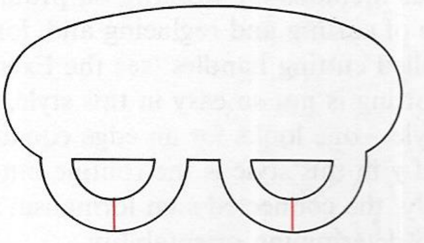


and here is another picture, obtained from the first by cutting along CD and retracting the arms.

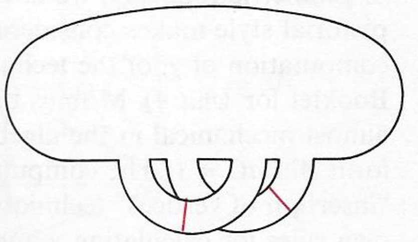


This picture supplies an edge equation $xyxz = 1$ for the surface, and so demonstrates non-orientability.

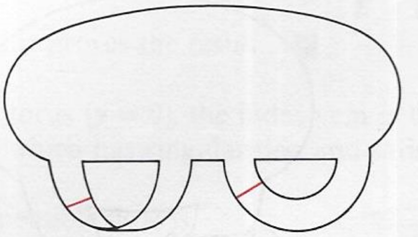
SAQ 9 More fun is to be had with the next six surfaces, which represent the possible ways in which two strips can be fastened to a disc. Are they all topologically distinct? Work through at least some of them if you want to revise the pictorial style of analysis. Cut along the internal edges to compute χ, and to determine orientability in difficult cases; use coloured pens to calculate β.



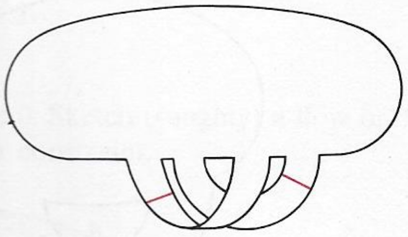
S_3



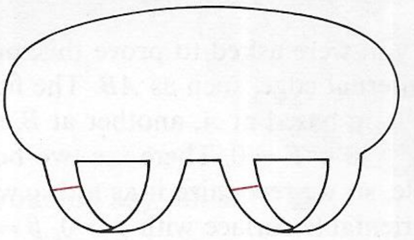
S_4



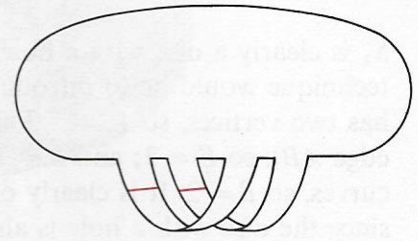
S_5



S_6



S_7

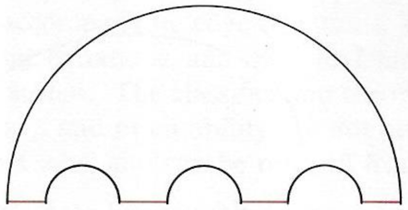


S_8

SOLUTION TO SAQ 9

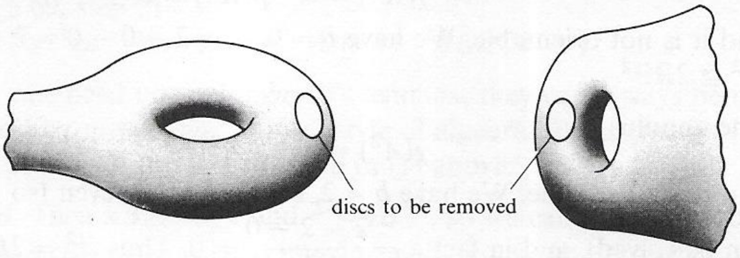
	χ	β	Orientable
S_3	-1	3	✓
S_4	-1	1	✓
S_5	-1	2	×
S_6	-1	1	×
S_7	-1	1	×
S_8	-1	2	×

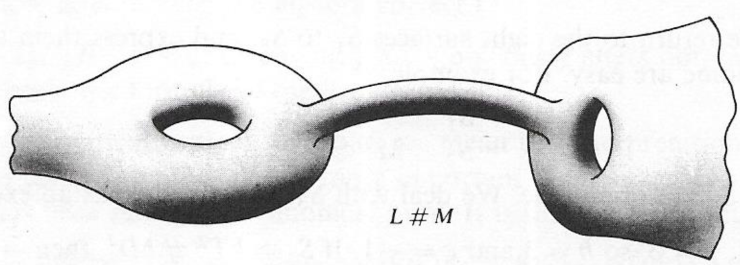
To verify our answers, observe that cutting each surface along the two marked internal edges produces a polygon like this each time, with various edge identifications to be made along the straight edges.



In each case, the edges are identified in pairs, so we find $V = 4$, $E = 6$, $F = 1$, $\chi = -1$. Orientability is determined immediately from the simpler figures S_1 and S_2, so the problem reduces to finding β by using coloured pens. The result is that two pairs are homeomorphic: $S_5 \cong S_8$, and $S_6 \cong S_7$. ■

We want to interpret our results in the geometric style, i.e. in terms of connected sums, because that is the simplest language with which to describe a surface; so let us recall how that goes. We denote a sphere by S^2, a disc by D^2, a torus by T^2, and a real projective plane by $\mathbb{RP}^2$. We write the connected sum of two surfaces L and M as $L \# M$. It is formed by removing a disc from each of L and M and then glueing in a handle (a cylinder) along the boundary.





Since $\chi(L - D^2) = \chi(L) - 1$ and $\chi(M - D^2) = \chi(M) - 1$, and glueing in the handle does not change the Euler characteristic, we have

$$\chi(L \# M) = \chi(L) + \chi(M) - 2.$$

We also write $L \# L \# \ldots \# L$ (n times) as nL.

The following calculations are routine: $\chi(S^2) = 2$, $\chi(D^2) = 1$, so $\chi(S^2 \# D^2) = 1$, and more generally $\chi(S^2 \# nD^2) = 2 - n$. $\chi(S^2) = 2$, $\chi(T^2) = 0$, so $\chi(S^2 \# T^2) = 0$, and more generally $\chi(S^2 \# kT^2) = 2 - 2k$. Similarly, $\chi(S^2 \# \mathbb{RP}^2) = 1$ and $\chi(S^2 \# m\mathbb{RP}^2) = 2 - m$. A picture makes clear that, if S is any surface, then $S^2 \# S \cong S$, so we may virtually suppress S^2 from the expression of a general surface as a connected sum. It is fairly easy to see (and was proved in the Exercise Booklet for *Unit 4*) that $L \# M$ is orientable if and only if L and M are both separately orientable, and it is trivial that $\beta(L \# M) = \beta(L) + \beta(M)$. With this information, we can catalogue all surfaces, as follows. To produce a surface with a specified value of β, say $\beta = b$, form $S^2 \# bD^2$ ($\cong bD^2$). Since $\beta(T^2) = 0 = \beta(\mathbb{RP}^2)$, no connected sums of bD^2 with tori or projective planes can have a different value of b. If the surface is to be non-orientable, it must contain at least one $\mathbb{RP}^2$ (in its expression as a connected sum). So, to find an orientable surface with given $\chi = c < 2$ and $\beta = b$, the only possibility is $S = kT^2 \# bD^2$, for which $\chi(S) = (2 - 2k) + (2 - b) - 2 = 2 - 2k - b$. We must satisfy $2 - 2k - b = c$, and so $\dfrac{2 - b - c}{2} = k$, which is possible if and only if $b + c$ is even.

It is easier to find a non-orientable surface with specified $\chi = c < 2$ and $\beta = b$. For, in this case, we consider

$$S = m\mathbb{RP}^2 \# bD^2,$$

for which

$$\chi(S) = (2 - m) + (2 - b) - 2 = 2 - m - b.$$

Now we must satisfy $2 - m - b = c$, which is easy: $m = 2 - b - c$, and there is no parity constraint on c.

It may be worthwhile pausing to consider some well-known surfaces which may not be so familiar as connected sums. For example, the Möbius band (M^2), the Klein bottle (K^2) and the cylinder or annulus (A^2). Let us take each in turn, and use the classification theorem to express them as connected sums.

(i) The Möbius band:

$$\chi(M^2) = 0, \quad \beta(M^2) = 1,$$

and it is certainly not orientable, so in the terms of our above argument, we have $b = 1$, $m = 2 - 1 - 0$ and $M^2 \cong \mathbb{RP}^2 \# D^2$.

(ii) The Klein bottle:

$$\chi(K^2) = 0, \quad \beta(K^2) = 0,$$

and it is not orientable. We have $b = 0$, $m = 2 - 0 - 0 = 2$ and $K^2 \cong 2\mathbb{RP}^2$.

(iii) The annulus:

$$\chi(A^2) = 0, \quad \beta(A^2) = 2,$$

and it is orientable. We have $b = 2$, $c = 0$, $b + c$ is even (so our problem can be solved), and in fact $k = \dfrac{2 - 2 - 0}{2} = 0$. Thus $A^2 = 2D^2$.

Now we return to the eight surfaces S_1 to S_8, and express them as connected sums. Some are easy. For example,

$$S_1 \cong A^2 \cong 2D^2,$$
$$S_2 \cong M^2 \cong \mathbb{RP}^2 \# D^2.$$

S_3 and S_4 are orientable. We deal with S_3, and leave S_4 as an exercise. For S_3, $\chi = -1$, $\beta = 3$, so $b = 3$ and $c = -1$. If $S_3 \cong kT^2 \# bD^2$, then $\dfrac{2 - 3 + 1}{2} = k$, i.e. $k = 0$, and $S_3 \cong 3D^2$.

S_5 ($\cong S_8$) and S_6 ($\cong S_7$) are non-orientable. Again we deal with one and leave the other as an exercise. Consider S_5: if $S_5 \cong m\mathbb{RP}^2 \# bD^2$, then $b = \beta(S_5) = 2$, and $m = 2 - b - c = 2 - 2 + 1 = 1$, so $S_5 \cong \mathbb{RP}^2 \# 2D^2$. This means it is homeomorphic to a Möbius band with a hole in it, a fact more evident in the S_5 picture than in the S_8 one.

SAQ 10 (i) Show $S_4 \cong T^2 \# D^2$.

(ii) Show $S_6 \cong 2\mathbb{RP}^2 \# D^2$.

SOLUTION TO SAQ 10 (i) It is important to realize that S_4 is not, like S_3, a disc with two holes, even though the pictures presented look rather alike. Because the two "strips" interweave, S_4 has *one* boundary curve, and $\beta = 1$.

(ii) S_6 is not orientable, because of the twist in the strip starting at the left. As in S_4, $\beta = 1$ and $\chi = -1$; S_4 is a Klein bottle with a hole, or $2\mathbb{RP}^2 \# D^2$. ■

It is time to return to the algebraic style. Numerous exercises in symbol manipulation have been provided in the relevant Exercise Booklet, and you should certainly do a few for practice, but it is easy to get lost in the web of letters; do not do too many. Here are examples of the basic methods.

EXAMPLE 8 Given $abacbc = 1$ to put into canonical form, eliminate a by writing $ba = x$.

Solution

$$\begin{aligned} abacbc = 1 \quad &\longrightarrow \quad axcbc = 1;\ ba = x \\ &\longrightarrow \quad b^{-1}xxcbc = 1. \end{aligned}$$

Repeating this trick on the cbc part re-writes the edge equation as $b^{-1}xxb^{-1}yy = 1$, and one eventually obtains the edge equation as $zzxxyy = 1$. □

A more systematic technique is to use Lemmas 1–2 of **B** Chapter 2.

For reference: they assert

1. $AxBCxD = 1 \longrightarrow AxCxB^{-1}D = 1$ and
   $\longrightarrow AC^{-1}xBxD = 1$
2. $AxBCx^{-1}D = 1 \longrightarrow AxCBx^{-1}D = 1$
   $ABxCx^{-1}D = 1 \longrightarrow AxCx^{-1}BD = 1$

Note also

3. (**B** page 69, step 2)
   $ABsCtDs^{-1}Et^{-1}F = 1 \longrightarrow Asts^{-1}t^{-1}BEDCF = 1$

Of course, one need not *remember* **B**'s lemmas; they can always be recovered (in emergencies) by use of the "simple" style of algebra. As an example, we shall show how to obtain the first formula in (1) above.

Let $y = xB$. Then $xBy^{-1} = 1$ and $x = yB^{-1}$, so we can replace xB by y and x by yB^{-1} in $AxBCxD = 1$. (Just imagine x, B and so on are matrices, if you want to know how to keep the algebra correct.)

This gives $AyCyB^{-1}D = 1$. Changing y for x obviously alters nothing, so we do this, and obtain the formula of the lemma.

It is the mathematical content, by which we mean the interpretation of an edge equation in canonical form, that is more important. So assume that a given equation has been reduced to canonical form. It is then either of this form

$$(1) \quad a_1b_1a_1^{-1}b_1^{-1}\ldots a_pb_pa_p^{-1}b_p^{-1}c_1d_1c_1^{-1}c_2d_2c_2^{-1}\ldots c_kd_kc_k^{-1} = 1,$$

or this

$$(2) \quad a_1a_1a_2a_2\ldots a_qa_qc_1d_1c_1^{-1}\ldots c_kd_kc_k^{-1} = 1.$$

The first form describes a p-holed torus with k holes, the second q copies of $\mathbb{RP}^2$ with k holes. We can use these examples to practise the technique of inserting vertices.

SAQ 11 Insert vertices in (1) to compute the Euler characteristic of the corresponding surface.

SOLUTION TO SAQ 11 Consider the start:

$$a_1b_1a_1^{-1}b_1^{-1}\ldots = 1,$$
$$(a_1b_1) = (b_1a_1^{-1}) = (a_1^{-1}b_1^{-1}) = (c_k^{-1}a_1),$$

so the equation begins

$$Pa_1Pb_1Pa_1^{-1}Pb_1^{-1}P\ldots = 1, \quad P = (a_1b_1).$$

Continuing the sequence, we find

$$Pa_1Pb_1Pa_1^{-1}Pb_1^{-1}P\ldots Pa_p^{-1}Pb_p^{-1}Pc_1Q_1d_1Q_1c_1^{-1}P\ldots Pc_kQ_kd_kQ_kc_k^{-1}P = 1.$$

Thus $V = k + 1$, $E = 2(p + k)$, $F = 1$, so $\chi = 2 - 2p - k$. Evidently $d_1, \ldots, d_k$ are boundary edges, so $\beta = k$. It is orientable, since every repeated edge occurs once as x, once as x^{-1}. When we compare this with $pT^2 \# kD^2$, we find $\chi(pT^2 \# kD^2) = 2 - 2p - k$, $\beta(pT^2 \# kD^2) = k$, and since this surface too is orientable, we deduce, by the fundamental theorem, that equation (1) describes a p-holed torus with k holes, as claimed. ■

SAQ 12 Analyse equation (2) in the same way.

SOLUTION TO SAQ 12 We find that the insertion of vertices yields

$$Pa_1Pa_1Pa_2Pa_2P\ \ldots\ Pa_qPa_qPc_1Q_1d_1Q_1c_1^{-1}\ \ldots\ Pc_kQ_kd_kQ_kc_k^{-1}P = 1.$$

Thus $V = 1 + k$, $E = q + 2k$, $F = 1$, $\chi = 2 - q - k$. Also, the surface is non-orientable, because the edge a_1 occurs twice as a_1 (and so does each a_i). This agrees with $q\mathbb{RP}^2 \# kD^2$, so, by the classification theorem, (2) is an edge equation for $q\mathbb{RP}^2 \# kD^2$. ■

When analysing surfaces algebraically, it is not necessary to reduce the equation to canonical form. Consider, for example, these six surfaces (the numbers are deliberate).

$$\begin{aligned}
Q_3&: \ abcdc^{-1}ea^{-1}f = 1\\
Q_4&: \ abcda^{-1}ec^{-1}f = 1\\
Q_5&: \ abcdc^{-1}eaf = 1\\
Q_6&: \ abcdaec^{-1}f = 1\\
Q_7&: \ abcdceaf = 1\\
Q_8&: \ abcdaecf = 1
\end{aligned}$$

Of course, these are edge equations for the surfaces S_3 to S_8, respectively, but if you did not know that, how would you prove it? One strategy would be: insert vertices, compute χ and β, look for repeated edges $\dots x \dots x^{-1} \dots$ to determine orientability; likewise analyse each of S_0 to S_8 (as has already been done); then, finally, use the classification theorem to decide which equations describe which surfaces. Other strategies, such as rearranging equations or cutting and glueing surfaces until recognizable forms are obtained, are of course possible. We leave it as an exercise for you (that is, we do not give the solution) to see that each equation has in fact been obtained by starting at the top (or right) endpoint of the left-hand edge and working down (or left).

SAQ 13 Analyse each equation Q_3 to Q_8, and also reduce it to canonical form. Hence confirm our conclusions so far.

Our answers are given at the end of the section. Work as many as you feel you need to do to revise these techniques, but leave yourself enough time to do the rest of your revision!

We hope that by now you have reviewed your application of the classification theorem, and how it can be used to analyse a surface presented in pictorial, algebraic or geometric terms. Moreover, in getting this far, you should have been able to practise most of the techniques that this course involves. For the rest of this revision section, we shall look at the other substantial topic dealt with in the course: covering spaces.

SOLUTION TO SAQ 13 Q_3 Inserting vertices gives

$$PaQbRcSdSc^{-1}ReQa^{-1}PfP = 1,$$

so $V = 4$, $E = 6$, $F = 1$; $\chi = -1$.

Boundaries are $QbReQ$, SdS, PfP; $\beta = 3$.

The repeated edges occur as $\dots a \dots a^{-1} \dots, \dots c \dots c^{-1} \dots$, so the surface is orientable.

The surface has the same characteristics as S_3. To obtain a canonical form for the edge equation, write it as:

$$\begin{aligned}
bcdc^{-1}ea^{-1}fa = 1 \ &\longrightarrow \ becdc^{-1}a^{-1}fa = 1\\
&\longrightarrow \ xcdc^{-1}a^{-1}fa = 1 \quad (\text{writing } x = be)\\
&\longrightarrow \ yxy^{-1}cdc^{-1}zfz^{-1} = 1,
\end{aligned}$$

the canonical form for a disc with two holes.

Q_4 Inserting vertices gives

$$PaQbRcSdQa^{-1}PeSc^{-1}RfP = 1,$$

so $V = 4$, $E = 6$, $F = 1$; $\chi = -1$.

Boundary is $QbRfPeSdQ$; $\beta = 1$.

The repeated edges occur as $\dots a \dots a^{-1} \dots, \dots c \dots c^{-1} \dots$, so the surface is orientable.

The surface has the same characteristics as S_4. To obtain a canonical equation for the edge equation, write it as:

$$\begin{aligned} abcda^{-1}ec^{-1}f=1 &\longrightarrow aca^{-1}c^{-1}edbf=1\\ &\longrightarrow aca^{-1}c^{-1}xyx^{-1}=1, \end{aligned}$$

the canonical form for a torus with a hole.

Q_5 Inserting vertices gives

$$PaQbRcSdSc^{-1}RePaQfP=1,$$

so $V=4$, $E=6$, $F=1$; $\chi=-1$.

Boundaries are $QbRePf^{-1}Q$, SdS; $\beta=2$.

The edge a occurs as $\dots a \dots a \dots$, so the surface is not orientable.

It has the same characteristics as S_5. To obtain a canonical form, write the edge equation as

$$\begin{aligned} abcdc^{-1}eaf=1 &\longrightarrow aae^{-1}cd^{-1}c^{-1}b^{-1}f=1,\\ aae^{-1}cd^{-1}c^{-1}b^{-1}f=1 &\longrightarrow aacd^{-1}c^{-1}e^{-1}b^{-1}f=1,\\ aacd^{-1}c^{-1}e^{-1}b^{-1}f=1 &\longrightarrow aacd^{-1}c^{-1}xyx^{-1}=1, \end{aligned}$$

the canonical form for a projective plane with two holes.

Q_6 Inserting vertices gives

$$PaQbRcSdPaQeSc^{-1}RfP=1,$$

so $V=4$, $E=6$, $F=1$; $\chi=-1$.

Boundary is $QbRfPd^{-1}Se^{-1}Q$; $\beta=1$.

Since a occurs as $\dots a \dots a \dots$, the surface is not orientable.

It has the same characteristics as S_6. To obtain a canonical form for the edge equation, write it as

$$\begin{aligned} abcdaec^{-1}f=1 &\longrightarrow aad^{-1}c^{-1}b^{-1}ec^{-1}f=1\\ &\longrightarrow aad^{-1}c^{-1}c^{-1}e^{-1}bf=1\\ &\longrightarrow aac^{-1}c^{-1}xyx^{-1}=1, \end{aligned}$$

for a Klein bottle with a hole.

Q_7 Since S_7 and S_6 are homeomorphic, we shall be briefer here. Inserting vertices gives

$$PaQbRcSdRcSePaQfP=1,$$

so $V=4$, $E=6$, $F=1$; $\chi=-1$.

Boundary is $QbRd^{-1}SePf^{-1}Q$; $\beta=1$.

Since a occurs as $\dots a \dots a \dots$, the surface is not orientable, and it therefore has the same characteristics as S_7 (and S_6).

We omit the reduction of the edge equation to canonical form.

Q_8 Since $S_8 \cong S_5$, we can again be brief. Inserting vertices gives

$$PaQbRcSdPaQeRcSfP=1,$$

so $V=4$, $E=6$, $F=1$; $\chi=-1$.

Boundaries are $QbRe^{-1}Q$, $SdPf^{-1}S$; $\beta=2$.

Since the edge a occurs as $\dots a \dots a \dots$, the surface is not orientable. It has the same characteristics as S_8 (and S_5).

We omit the reduction to canonical form. ■

Part II: Covering Surfaces

This topic fell into two pieces: complex conics, and more general covering surfaces. Within the story about conics, there were certain subdivisions, as follows. An irreducible complex quadratic equation $f(z, w) = 0$ defines a locus (the complex conic) as z runs over the plane of complex numbers, $\mathbb{C}$. Neither the locus nor $\mathbb{C}$ are compact, so these are not surfaces. We make $\mathbb{C}$ compact by adding the "point at infinity" to form $\mathbb{C} \cup \{\infty\}$: we must then also make the locus compact. There are two ways of doing this. The algebraic (adding points at infinity) way can fail to yield a surface, because pinch points may arise; but the geometric way always yields a surface, the so-called Riemann surface corresponding to the equation. The main theorem of *Unit 5* describes the result of algebraically compactifying the general conic, but it is perhaps more worthwhile to be sure you can compactify any given conic, for example the ones given immediately below. The result is always one of the following:

(i) a surface homeomorphic to the z-sphere;

(ii) a two-fold covering of the z-sphere by a surface homeomorphic to a sphere, and branched (as a covering) over two points;

(iii) a two-fold covering branched over two points but with a pinch point (this is the "usual" case).

Here, in no particular order, are a representative set of examples of conics, and one quadratic which is not a conic. Work as many of these examples as you feel like doing without quoting the theorems. Full solutions are given for convenience of reference (and not to pressure you to do the lot!). Recall that the only techniques involved are:

(i) replacing z by $\frac{1}{z} = \tilde{z}$ and w by $\frac{1}{w} = \tilde{w}$ when looking for "infinite" points;

(ii) discriminating between pinch points and branch points.

SAQ 14

(i) $2z^2 - zw - w^2 + 3z + 3w - 2 = 0$

(ii) $z^2 + 4zw + w^2 + 3z = 0$

(iii) $2zw + w^2 + w + z = 0$

(iv) $z^2 + w^2 + w = 0$

(v) $4z^2 + 4zw + w^2 + 2z + 1 = 0$

(vi) $zw + 3 = 0$

(vii) $z^2 + 5zw + w = 0$

SOLUTION TO SAQ 14

(i) This is not a conic. Writing it as

$$w^2 + (z - 3)w + (-2z^2 - 3z + 2) = 0$$

and solving it for w by the usual formula, we find

$$w = \frac{-(z-3) \pm \sqrt{((3z+1)^2)}}{2},$$

so $(w - z - 2)(2z + w - 1) = 0$, and the equation is reducible.

(ii) $w^2 + 4zw + z^2 + 3z = 0$

$$\begin{aligned} &\Rightarrow w = -2z \pm [4z^2 - z^2 - 3z]^{1/2} \\ &\quad\;\; = -2z \pm \sqrt{3}\sqrt{z(z-1)}, \end{aligned}$$

so $z = 0$, $w = 0$, and $z = 1$, $w = -2$, are branch points. Now we look for infinite points.

$(z, w) = (\infty, w)$, $w \neq \infty$, lies on the locus
$\iff \quad (\tilde{z}, w) = (0, w) \quad$ lies on $\quad 1 + 4\tilde{z}w + \tilde{z}^2w^2 + 3\tilde{z} = 0,$
which it does not.

$(z, w) = (z, \infty)$, $z \neq \infty$, lies on the locus
$\iff \quad (z, \tilde{w}) = (z, 0) \quad$ lies on $\quad z^2\tilde{w}^2 + 4z\tilde{w} + 1 + 3z\tilde{w}^2 = 0,$
which it does not.

$(z, w) = (\infty, \infty)$ lies on the locus
$\iff \quad (\tilde{z}, \tilde{w}) = (0, 0) \quad$ lies on $\quad \tilde{z}^2 + 4\tilde{z}\tilde{w} + (1 + 3\tilde{z})\tilde{w}^2 = 0,$
which it does, as a double point.

So (∞, ∞) is also an anomalous point. Since the points $(z, w) = (0, 0)$ and $(z, w) = (1, -2)$ are branch points, and the cover is a double cover, the third point cannot be a branch point, and must therefore be a pinch point. (This argument is standard and will not be repeated in the subsequent answers.)

(iii) $w^2 + (2z+1)w + z = 0$

$$\Longrightarrow \quad w = \frac{-(2z+1) \pm [4z^2 + 4z + 1 - 4z]^{1/2}}{2}$$

$$\Longrightarrow \quad w = -z - \tfrac{1}{2} \pm (z^2 + \tfrac{1}{4})^{1/2},$$

so $(z,w) = \left(\frac{i}{2}, \frac{-(1+i)}{2}\right)$ and $\left(-\frac{i}{2}, \frac{i-1}{2}\right)$ are branch points.

$(z,w) = (\infty, w)$, $w \neq \infty$ lies on the locus
$\Longleftrightarrow \quad (\tilde{z}, w) = (0, w)$ lies on $w^2\tilde{z} + (2+\tilde{z})w + 1 = 0$,
which is true for $w = -\frac{1}{2}$. So $(z,w) = (\infty, -\frac{1}{2})$ does lie on the locus.

$(z,w) = (z, \infty)$, $z \neq \infty$, lies on the locus
$\Longleftrightarrow \quad (z, \tilde{w}) = (z, 0)$ lies on $1 + (2z+1)\tilde{w} + z\tilde{w}^2 = 0$,
which it does not.

$(z,w) = (\infty, \infty)$ lies on the locus
$\Longleftrightarrow \quad (\tilde{z}, \tilde{w}) = (0,0)$ lies on $\tilde{z} + (2+\tilde{z})\tilde{w} + \tilde{w}^2 = 0$,
which it does.

The infinite points on the locus are $(\infty, -\frac{1}{2})$ and (∞, ∞), and the locus is a sphere. It is a double cover of the z-sphere, branched over $z = \frac{1}{2}$ and $z = -\frac{1}{2}$.

(iv) $w^2 + w + z^2 = 0$

$$\Longrightarrow \quad w = \frac{-1 \pm [1 - 4z^2]^{1/2}}{2},$$

so $(z,w) = (\frac{1}{2}, -\frac{1}{2})$ and $(-\frac{1}{2}, -\frac{1}{2})$ are branch points.

$(z,w) = (\infty, w)$, $w \neq \infty$, lies on the locus
$\Longleftrightarrow \quad (\tilde{z}, w) = (0, w)$ lies on $(w^2 + w)\tilde{z}^2 + 1 = 0$,
which it does not.

$(z,w) = (z, \infty)$, $z \neq \infty$, lies on the locus
$\Longleftrightarrow \quad (z, \tilde{w}) = (z, 0)$ lies on $1 + \tilde{w} + z^2\tilde{w}^2 = 0$,
which it does not.

$(z,w) = (\infty, \infty)$ lies on the locus
$\Longleftrightarrow \quad (\tilde{z}, \tilde{w}) = (0,0)$ lies on $(1 + \tilde{w})\tilde{z}^2 + \tilde{w}^2 = 0$,
which it does, as a double point.

The locus is a sphere, double covering the z-sphere, branched over $z = \frac{1}{2}$, $z = -\frac{1}{2}$, and with (∞, ∞) as a pinch point.

(v) $w^2 + 4zw + 4z^2 + 2z + 1 = 0$

$$\Longrightarrow \quad w = -2z \pm [4z^2 - 4z^2 - 2z - 1]^{1/2} = 0,$$

so $(z,w) = (-\frac{1}{2}, 1)$ is a branch point.

Since the covering must have two branch points, the other one must be at infinity.

$(z,w) = (\infty, w)$, $w \neq \infty$, lies on the locus
$\Longleftrightarrow \quad (\tilde{z}, w) = (0, w)$ lies on $w^2\tilde{z}^2 + 4\tilde{z}w + 4 + 2\tilde{z} + \tilde{z}^2 = 0$,
which it does not.

(z, ∞), $z \neq \infty$, lies on the locus
$\Longleftrightarrow \quad (z, \tilde{w}) = (z, 0)$ lies on $4z^2\tilde{w}^2 + 4z\tilde{w} + 1 + 2z\tilde{w}^2 + \tilde{w}^2 = 0$,
which it does not.

$(z,w) = (\infty, \infty)$ lies on the locus
$\Longleftrightarrow \quad (\tilde{z}, \tilde{w}) = (0,0)$ lies on $\tilde{z}^2 + 4\tilde{z}\tilde{w} + (4 + 2\tilde{z} + \tilde{z}^2)\tilde{w}^2 = 0$,
which it does.

The locus is a sphere, double covering the z-sphere, and branched at $(z,w) = (-\frac{1}{2}, 1)$ and (∞, ∞).

(vi), (vii) There is no w^2 term, so the locus is homeomorphic to the z-sphere. It is not branched, being a "1-fold cover".) ■

The theory of complex conics is a part of a general theory of covering surfaces. Roughly speaking, a covering is a continuous map $\tilde{S} \xrightarrow{\pi} S$ between two surfaces in which every point of S has the same number of pre-images. This definition must be modified to allow for branched coverings, and *Unit 6* described how this is done. We described three ways in which covering surfaces can be produced:

(i) *geometrically*, by exploiting a symmetry of $\tilde{S}$ to find S;

(ii) *combinatorially*, by using the edge equation for S to find an edge equation for $\tilde{S}$;

(iii) *algebraically*, by interpreting a locus in (z, w) as a Riemann surface over the z-sphere (called, in *Unit 6*, the geometric method for compactifying a locus).

The first two can be illustrated with the same example, which is the best example to remember: $\tilde{S} = S^2$, $S = \mathbb{RP}^2$, and the map sends $P \in S^2$ to the projective point defined by P and its antipodal point.

The third approach has been amply illustrated with conic sections, but extends to other loci. The easiest to understand are $w^2 =$ (polynomial in z) and $w^n = z$.

The topology comes in via the classification theorem. We seek to relate $\chi(\tilde{S})$ to $\chi(S)$, $\beta(\tilde{S})$ to $\beta(S)$, and the orientability of $\tilde{S}$ to that of S. Assume first that the covering is n-fold and unbranched. Then the Euler characteristics are neatly related:

$$\chi(\tilde{S}) = n\chi(S).$$

(Why? Can you remember the proof?) The relationship between the βs is not so simple:

$$\beta(S) \le \beta(\tilde{S}) \le n\beta(S).$$

The most useful thing to remember is that if S has no boundary then $\tilde{S}$ can have no boundary.

Finally, orientation. The example $S^2 \longrightarrow \mathbb{RP}^2$ almost says it all; everything can happen *except* non-orientable covering orientable.

The relationship $\chi(\tilde{S}) = n\chi(S)$ for n-fold unbranched coverings has some pleasing consequences. Some questions we shall not answer here are:

What can n and $\tilde{S}$ be if $S = S^2$ $(\chi(S^2) = 2)$?
What if $S = T^2$ $(\chi(T^2) = 0)$?
Can you ever have $\chi(\tilde{S}) > \chi(S)$?

When coverings are branched, the discussion of β and orientation is unaltered, but the Euler characteristic formula must be modified. It becomes

$$\chi(\tilde{S}) = n\chi(S) - \delta,$$

where δ is the total branch point order,

$$\delta = \sum_{P \in S} (n - \text{number of points in } \pi^{-1}(P)).$$

(Notice that this sum is 0 for all but the finitely many points P in S over which the surface is branched.)

The connection with Riemann surfaces, defined as coverings of the Riemann sphere S^2, is that all but the trivial 1-fold cover are branched, orientable, and without boundary. Consequently, they are classified by their genus. However, the surface S need not be S^2 when a general covering is considered.

Let us turn to techniques. Perhaps rather arbitrarily we select two, one combinatorial and the other algebraic. Combinatorially, suppose you are given $aba^{-1}b^{-1}cdc^{-1}d^{-1} = 1$, the canonical equation of a double torus. A double cover is given by the surface $\tilde{S}$, with edge equations

$$\begin{aligned} a_1b_1a_2^{-1}b_2^{-1}c_1d_1c_2^{-1}d_2^{-1} &= 1,\\ a_2b_2a_1^{-1}b_1^{-1}c_2d_2c_1^{-1}d_1^{-1} &= 1. \end{aligned}$$

What surface is it? It would be possible to answer this question by finding the new canonical equation, but that is very tedious. The surface is certainly orientable and without boundary (why?), and so is classified by its genus. To find that, the simplest technique is to insert vertices. When we do this, we find we can write

$$P_1a_1P_2b_1P_3a_2^{-1}P_4b_2^{-1}P_5c_1P_6d_1P_7c_2^{-1}P_8d_2^{-1}P_1 = 1,$$

but the second equation forces the identifications $P_1 = P_3 = P_5 = P_7$, $P_2 = P_4 = P_6 = P_8$. So the new surface has $V = 2$, $E = 8$, $F = 2$, $\chi = -4$, and is therefore a 3-holed torus. One can reach the same conclusion another way. We know that the base surface has only one vertex (see above, p. 53) and if, as is the case, this vertex lifts to two, then the covering surface is not branched. Indeed, with this particular covering, the only way $\tilde{S}$ could be the k-fold torus would be if there were $2k - 6$ branch points in the covering, corresponding to $2 - 2k = 2(-2) - (2k - 6)$, $\delta = 2k - 6$. Since P is a single point, this branched covering is impossible, and so one can deduce without any calculation that edge equation $(*)$ describes a two-fold unbranched cover of $2T^2$ by $3T^2$. We leave it to you to find a geometric realization of this covering. (Look for a rotation of a suitable $3T^2$ through 180°.)

Suppose we wanted a three-fold covering of the double torus by the quadruple torus (sometimes denoted $4T^2 \xrightarrow{3} 2T^2$) corresponding to $-6 = \chi(4T^2) = 3(-2) = 3\chi(2T^2)$. $2T^2$ is described by the edge equation

$$PaPbPa^{-1}Pb^{-1}PcPdPc^{-1}Pd^{-1}P = 1.$$

Consider the surface $\tilde{S}$ defined by

$$\begin{aligned} &P_1a_1P_2b_1P_3a_2^{-1}P_4b_2^{-1}P_5c_1P_6d_1P_7c_2^{-1}P_8d_2^{-1}P_1 = 1,\\ &a_2b_2a_3^{-1}P_9b_3^{-1}c_2d_2c_3^{-1}P_{10}d_3^{-1} = 1,\\ &a_3b_3a_1^{-1}b_1^{-1}c_3d_3c_1^{-1}d_1^{-1} = 1. \end{aligned}$$

To insert any more vertices is to invite equalities. In fact, we find $P_1 = P_3 = P_5 = P_7$, $P_2 = P_6 = P_9 = P_{10}$, $P_4 = P_8$. The surface is unbranched, and is therefore the one we seek. (Had we found $P_1 = P_2 = P_3 = P_4 = \ldots = P_{10}$, we should have had $\chi(\tilde{S}) = 3 \times (-2) - 2 = -8$, making $\tilde{S} = 5T^2$.) The geometrical realization involves finding a representation of $4T^2$ with a 120° symmetry, but of course the edge equations alone guarantee the existence of the covering.

It remains for us to wish you well in the examination and in your future study, and to express the hope that you have enjoyed the course.

0 335 14021 1